BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

LAS INVERSAS GENERALIZADAS DE LAS MATRICES

Tesis para obtener el título de:
Licenciado en Matemáticas

por

Sergio Atayan García Balán

asesorada por

Dr. Slavisa Djordjevic

Puebla Pue.

Junio de 2011

A Rogelio, quien ya está leyendo (y seguramente ampliando), el libro de las demostraciones excelsas.

A Naomi, que está por llegar.

Llegué a mi casa después de haberme visto entre un trío de hombres vestidos y en estado reproductivo. Abrí la gruesa ventana que le impide al conejo salir de su prisión y lo tomé entre mis brazos. Después de varios intentos de escapar, logré que se tranquilizara y con mis manos lo acaricié mientras una postura de padre responsable lo sostuvo con mis brazos y lo arrulló con mi voz junto a mi pecho.

Hoy el conejo murió o casi ha muerto. Yo casi lo maté pero me he resistido. A veces me imagino taladrando su pescuezo. Dejándolo apenas vivo pero con ganas de estar muerto. Tornándolo iracundo y con ganas de venganza. Con ganas de que un día cualquiera, mientras escribo, me odie como lo odio ahora y casi me apuñale como casi lo apuñalo hoy.

El conejo vive detrás de la ventana. Lo miro tirado junto a su traste de agua lamiéndose las orejas y haciendo cosas de conejo, ruidos de conejo. Lo miro y siento que piensa como conejo, cada brinco lo acerca más a su padre. A veces siento lástima por él y a veces simplemente me da risa. Mientras más lo contemplo más me doy cuenta de su realidad, aunque en el fondo sé que esa ventana siempre ha sido un espejo.[0]

[0] El conejo, Oscar Bacerott.

Agradecimientos

Gracias a toda mi familia por lo que me ha dado, en especial a mis padres que con su ejemplo me han enseñado a amar, a pensar, a vivir; a mi hermano por darme mis primeros golpes de la vida y que ahora no hace más que protegerme; a Yáscara y a Andy a quienes quiero mucho.

Gracias a todos mis amigos, sin ellos se hubiera complicado bastante mi andar. En especial a Vianey, Nelson y René quienes me ayudaron infinitamente con este trabajo.

Agradezco a los profesores de nuestra facultad. Ellos me fueron mostrando en estos años las responsabilidades que conlleva ser un matemático en un país como México. En especial a Manuel Ibarra por brindarme su sincera amistad; y a Slavisa Djordjevic por toda la confianza que ha tenido en mí y el apoyo que me ha dado.

Un agradecimiento especial a la profesora Guadalupe Raggi, al profesor Alberto Escamilla y al profesor Jacobo Oliveros por aceptar ser mis sinodales y mejorar así, en gran medida, esta tesis.

Prefacio

Un tema importante en el Álgebra Lineal, es el estudio de las matrices. Parte de su relevancia radica en que son auxiliares en la resolución de sistemas de ecuaciones lineales, los cuales se presentan con mucha frecuencia en diversas áreas científicas, tanto naturales como sociales, en la estadística, computación, etc. En particular, las matrices invertibles cumplen ciertas propiedades que son de utilidad, es por ello que dada una matriz, nos preguntamos si ésta es invertible. Desafortunadamente, en muchos casos no lo es.

La teoría de los inversos generalizados, entre otras cosas, se ocupa de las matrices que no son invertibles, definiendo para ellas, ciertas matrices, que llamaremos inversas generalizadas, y que poseen cierta similitud con la inversa usual (en caso de existir); esto claro está, con el objetivo de rescatar algunas propiedades que se pueden pensar perdidas al no contar con una matriz invertible.

En este trabajo presentaremos a las inversas generalizadas. Empezaremos por su existencia y construcción. Después, nos ocuparemos de encontrar caracterizaciones de varias de ellas y, finalmente veremos algunas de sus aplicaciones.

Es importante mencionar que la teoría de los inversos generalizados, a pesar de tener poco más de un siglo de existencia, continua siendo un campo grande de investigación que arroja resultados de forma cotidiana. Es precisamente por eso, que creemos importante el estudio de la misma.

Sergio Atayan García Balán
Facultad de Ciencias Físico Matemáticas,
Benemérita Universidad Autónoma de Puebla.
Junio de 2011

Índice general

Capítulo 1

Preliminares

Minino de Cheshire, ¿podrías decirme, por favor, qué camino debo seguir para salir de aquí? -Esto depende en gran parte del sitio al que quieras llegar -dijo el Gato. -No me importa mucho el sitio... -dijo Alicia. -Entonces tampoco importa mucho el camino que tomes -dijo el Gato. -... siempre que llegue a alguna parte -añadió Alicia como explicación. -¡Oh, siempre llegarás a alguna parte -aseguró el Gato-, si caminas lo suficiente!.[1]

Comenzaremos el capítulo con una introducción a las inversas generalizadas. En las secciones restantes, daremos a conocer la notación que emplearemos así como las defíniciones y los resultados básicos que se usarán en los siguientes capítulos.

1.1. Introducción

Una matriz tiene una inversa sólo si es cuadrada y, además, es no singular, o en otras palabras, si sus columnas (o filas) son linealmente independientes. A mediados del siglo pasado se comenzó a sentir la necesidad, en numerosas áreas de las matemáticas, por algún tipo de inversa parcial de una matriz

[1] Alicia en el país de las maravillas, Lewis Carroll.

que sea singular o rectangular. Por la inversa generalizada de una matriz A entenderemos una matriz X asociada de alguna manera con A y tal que:

- exista para una clase de matrices mayor que la clase de matrices no singulares;
- tenga algunas de las propiedades de la inversa usual; y
- se reduzca a la inversa usual cuando A sea no singular.

El concepto de "una inversa generalizada", parece haber sido mencionada primero (por escrito), por Fredholm en 1903, donde dio una inversa generalizada particular (llamada por él "pseudo inversa"), de un operador integral. La clase de todas las pseudo inversas fue caracterizada por Hurwitz en 1912. Los inversos generalizados de operadores diferenciales, implícitos en el trabajo de Hilbert sobre funciones de Green generalizadas en 1904, fue estudiado por numerosos autores, entre ellos Myller en 1906, Westfall en 1909.

De esta manera, los inversos generalizados de operadores diferenciales e integrales antecedieron a las inversas generalizadas de matrices, cuya existencia fue notada primero por E. H. Moore alrededor de 1906, quien definió una inversa única (llamada por él "la recíproca general") para cada matriz finita (cuadrada o rectangular). Sin embargo, el trabajo de Moore fue cayendo en el olvido durante algunos años debido, entre otras cosas, a que usaba notación muy engorrosa, lo cual hacía su trabajo ininteligible a la mayoría. Fue hasta 1955 que R. Penrose reavivó el interés en la materia con su trabajo. Debido a lo cual el estudio de los inversos generalizados, ha florecido a partir de la segunda mitad del siglo pasado, demostrando que tiene aplicaciones en muchas partes. Una nota histórica más completa y con referencias, puede encontrarse en [1].

1.2. Escalares y Vectores

Denotaremos a los *escalares* con letras minúsculas: x, y, $\lambda, \ldots$. Trabajaremos principalmente en $\mathbb{C}$, *el campo de los números complejos*. $\mathbb{R}$ denotará al *campo de los números reales* y $\mathbb{F}$ denotará algún campo arbitrario.

A los *vectores* los denotaremos con letras en negritas $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}, \ldots$. Los espacios vectoriales en los que trabajaremos (salvo que se diga lo contrario),

son de dimensión finita. El espacio vectorial n-dimensional sobre un campo $\mathbb{F}$ es denotado por $\mathbb{F}^n$, en particular $\mathbb{C}^n$ [$\mathbb{R}^n$] denota el espacio vectorial complejo [real] n-dimensional.
Un vector $\mathbf{x} \in \mathbb{F}^n$ se escribirá en forma de columna

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = (x_i), \quad i \in \overline{1,n}\,, \quad x_i \in \mathbb{F}.$$

El vector n-dimensional $\mathbf{e}_i$ con componentes

$$\delta_{ij} = \begin{cases} 1, & \text{si} \quad i = j, \\ 0, & \text{de otra manera,} \end{cases}$$

es llamado el *i-ésimo vector unitario* de $\mathbb{F}^n$. El conjunto $\mathcal{E}_n$ de vectores unitarios

$$\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\} \tag{1.2.1}$$

es llamado *base estándar* de $\mathbb{F}^n$.

La *suma* de dos conjuntos L, M en $\mathbb{C}^n$, denotada por $L + M$ se define como

$$L + M \;=\; \{\mathbf{y} + \mathbf{z} : \mathbf{y} \in L,\, \mathbf{z} \in M\}.$$

Si L y M son subespacios de $\mathbb{C}^n$, entonces $L + M$ es también un subespacio de $\mathbb{C}^n$. Si además $L \cap M = \{\mathbf{0}\}$, entonces $L + M$ es llamado la *suma directa* de L y M, denotada por $L \oplus M$. Dos subespacios L y M de $\mathbb{C}^n$ son llamados *complementarios* si

$$L \oplus M \;=\; \mathbb{C}^n. \tag{1.2.2}$$

Cuando esto sucede, cada $\mathbf{x} \in \mathbb{C}^n$ se puede expresar de forma única como una suma

$$\mathbf{x} = \mathbf{y} + \mathbf{z} \quad (\mathbf{y} \in L,\, \mathbf{z} \in M). \tag{1.2.3}$$

En este caso, diremos que $\mathbf{y}$ es la *proyección* de $\mathbf{x}$ sobre L a lo largo de M.

Sea V un espacio vectorial complejo. Un *producto interior* es una función: $V \times V \to \mathbb{C}$, denotada por $\langle \mathbf{x}, \mathbf{y} \rangle$, que satisface:

(I1) $\langle \alpha\mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \alpha\langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$ (linealidad);

(I2) $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$ (Simetría Hermitiana); y

(I3) $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$, $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ si y sólo si $\mathbf{x} = 0$ (positividad);

para todo $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ y $\alpha \in \mathbb{C}$.
El producto interno estándar en $\mathbb{C}^n$ es

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^* \mathbf{x} = \sum_{i=1}^{n} x_i \, \overline{y_i}, \tag{1.2.4}$$

para todo $\mathbf{x} = (x_i)$, $\mathbf{y} = (y_i)$ en $\mathbb{C}^n$.

Diremos que dos vectores son *ortogonales* si y sólo si su producto interior es igual a 0.

Sea V un espacio vectorial complejo. Una *norma* es una función: $V \to \mathbb{R}$, denotada por $\|\mathbf{x}\|$, que satisface:

(N1) $\|\mathbf{x}\| \geq 0$, $\|\mathbf{x}\| = 0$ si y sólo si $\mathbf{x} = 0$ (positividad);

(N2) $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$ (homogeneidad positiva);

(N3) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (desigualdad del triángulo);

para todo $\mathbf{x}, \mathbf{y} \in V$ y $\alpha \in \mathbb{C}$.
Salvo que se especifique lo contrario, si $\mathbf{x} \in \mathbb{C}^n$, entonces

$$\|x\| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}.$$

Lema 1.1. *Sean $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ vectores ortogonales, entonces*

$$\|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2.$$

Demostración. $\|\mathbf{x} + \mathbf{y}\|^2 = \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle$, luego dado que $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle = 0$ se tiene $\|\mathbf{x} + \mathbf{y}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$. $\square$

1.3. Transformaciones Lineales y Matrices

Una *matriz de $m \times n$ con elementos en un campo* $\mathbb{F}$ es un arreglo rectangular de la forma

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

donde cada a_{ij} $(1 \leq i \leq m,\ 1 \leq j \leq n)$ es un elemento de $\mathbb{F}$.

El conjunto de las matrices de $m \times n$ con elementos en $\mathbb{F}$ será denotado como $\mathbb{F}^{m\times n}$. En particular $\mathbb{C}^{m\times n}$ $(\mathbb{R}^{m\times n})$ denota el conjunto de las matrices con entradas complejas (reales).

Una matriz $A \in \mathbb{F}^{m\times n}$ será llamada *cuadrada* si $m = n$ y será llamada *rectangular* si $m \neq n$.

Las entradas de una matriz $A \in \mathbb{F}^{m\times n}$ serán denotadas como a_{ij} o $A[i,j]$.

Dada una matriz A será llamada:

- *diagonal* si $A[i,i] = 0$ para toda $i \neq j$,
- *triangular superior* si $A[i,i] = 0$ si $i > j$ y
- *triangular inferior* si $A[i,i] = 0$ si $i < j$.

Una matriz diagonal A de $m\times n$ será denotada como $A = diag(a_{11}, \ldots, a_{pp})$, donde $p = \min\{m,n\}$.

Dada una matriz $A \in \mathbb{C}^{m\times n}$, su:

- *transpuesta*, es la matriz A^T tal que $A^T[i,j] = A[j,i]$ para todo i,j.
- *conjugada transpuesta*, es la matriz $A^* \in \mathbb{C}^{n\times m}$ tal que $A^*[i,j] = \overline{A[j,i]}$ para todo i,j.

Dada una matriz $A \in \mathbb{C}^{m\times n}$, es:

- *hermitiana* (*simétrica*) si $A = A^*$ (si A es real, $A = A^T$);
- *normal* si $AA^* = A^*A$; y
- *unitaria* (*ortogonal*) si $A^* = A^{-1}$ (si A es real, $A^T = A^{-1}$).

Dados dos espacios vectoriales U, V sobre un campo $\mathbb{F}$ y una función $T : U \to V$, diremos que T es *lineal* o que es una *transformación lineal*, si $T(\alpha\mathbf{x} + \mathbf{y}) = \alpha T\mathbf{x} + T\mathbf{y}$, para cada $\alpha \in \mathbb{F}$ y cada $\mathbf{x}, \mathbf{y} \in U$. El conjunto de transformaciones lineales de U en V será denotado como $\mathcal{L}(U,V)$, este conjunto es un espacio vectorial con las operaciones $T_1 + T_2$ y αT definidas como

$$(T_1 + T_2)\mathbf{x} = T_1\mathbf{x} + T_2\mathbf{x}, \quad (\alpha T)\mathbf{x} = \alpha(T\mathbf{x}) \ \ \forall \mathbf{x} \in U.$$

El elemento nulo de $\mathcal{L}(U, V)$ es la transformación O, donde $O\mathbf{x} = \mathbf{0}$ para cada $\mathbf{x} \in U$. La transformación identidad $I_U \in \mathcal{L}(U, U)$ está definida como $I_U\mathbf{u} = \mathbf{u}, \ \forall \mathbf{u} \in U$. Denotaremos a la transformación identidad I_U sólo como I salvo que sea necesario distinguirla.

Sea $T \in \mathcal{L}(U, V)$. Para cada $\mathbf{u} \in U$, el vector $T\mathbf{u} \in V$ es llamado la *imagen* de $\mathbf{u}$ (bajo T). La *imagen* de T, denotada como $R(T)$, es el conjunto de todas las imágenes bajo T, es decir,

$$R(T) = \{\mathbf{v} \in V : \mathbf{v} = T\mathbf{u} \text{ para algún } \mathbf{u} \in U\}.$$

El *rango* de T, denotado como *rango*(T), es la dimensión de $R(T)$.

Para cualquier $\mathbf{v} \in R(T)$, la *imagen inversa* $T^{-1}(\mathbf{v})$ es el conjunto

$$T^{-1}(\mathbf{v}) = \{\mathbf{u} \in U : T\mathbf{u} = \mathbf{v}\}.$$

En particular, el *espacio nulo* de T, denotado como $N(T)$, es la imagen inversa del vector $\mathbf{0} \in V$,

$$N(T) = \{\mathbf{u} \in U : T\mathbf{u} = \mathbf{0}\} = T^{-1}(\mathbf{0}).$$

Una relación importante entre $R(T)$ y $N(T)$ conocida como el teorema de la dimensión, es la siguiente: si $T \in \mathcal{L}(U, V)$ (donde V es de dimensión finita), entonces

$$dim(N(T)) + dim(R(T)) = dim(V). \tag{1.3.1}$$

$T \in \mathcal{L}(U, V)$ es *inyectiva* si para cada $\mathbf{x}, \mathbf{y} \in U$ tales que $T\mathbf{x} = T\mathbf{y}$ entoces $\mathbf{x} = \mathbf{y}$, o equivalentemente si para cada $\mathbf{v} \in R(T)$ la imagen inverda $T^{-1}\mathbf{v}$ consta de un solo vector. T es *sobreyectiva* si $R(T) = V$. Si T es inyectiva y sobreyectiva entonces tiene una *inversa* $T^{-1} \in \mathcal{L}(V, U)$ tal que

$$T^{-1}T = I_U \quad \text{y} \quad TT^{-1} = I_V, \tag{1.3.2}$$

en este caso T es llamada *invertible* o *no singular*.

Dados

- una transformación lineal $A \in \mathcal{L}(\mathbb{C}^n, \mathbb{C}^m)$; y
- dos bases $\mathcal{U} = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ y $\mathcal{V} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ de $\mathbb{C}^m$ y $\mathbb{C}^n$ respectivamente;

la *representación matricial* de A respecto de las bases $\{\mathcal{U}, \mathcal{V}\}$ es la matriz de $m \times n$ $A_{\{\mathcal{U},\mathcal{V}\}} = [a_{ij}]$ determinada (únicamente) como

$$A\mathbf{v}_j = \sum_{\mathbf{i}=1}^{m} a_{ij}\mathbf{u}_i, \quad j \in \overline{1,n}. \tag{1.3.3}$$

Para cada par de bases $\{\mathcal{U}, \mathcal{V}\}$, el sistema de ecuaciones (1.3.3) nos da una correspondencia inyectiva y biyectiva entre el espacio de las transformaciones lineales $\mathcal{L}(\mathbb{C}^n, \mathbb{C}^m)$ y el espacio de las matrices $\mathbb{C}^{m\times n}$, permitiéndonos el uso del símbolo A como la transformación lineal $A : \mathbb{C}^n \to \mathbb{C}^m$, definida por el producto matricial, y como la representación matricial $A_{\{\mathcal{U},\mathcal{V}\}}$.

Si A es una transformación lineal de $\mathbb{C}^n$ en sí misma, y $\mathcal{V} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es una base de $\mathbb{C}^n$, entonces la representación matricial $A_{\{\mathcal{V},\mathcal{V}\}}$ será denotada simplimente como $A_{\{\mathcal{V}\}}$. $A_{\{\mathcal{V}\}} = [a_{ij}] \in \mathbb{C}^{n\times n}$ que satisface

$$A\mathbf{v}_j = \sum_{i=1}^{n} a_{ij}\mathbf{v}_i \quad j \in \overline{1,n}. \tag{1.3.4}$$

La *base estándar* de $\mathbb{C}^n$ es la base $\mathcal{E}_n$ que consiste de los n vectores unitarios

$$\mathcal{E}_n = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\} \quad \text{(ver 1.2.1)}.$$

Nótese que la representación matricial, respecto de las bases $\{\mathcal{E}_m, \mathcal{E}_n\}$, de la transformación lineal $A \in \mathcal{L}(\mathbb{C}^n, \mathbb{C}^m)$ coincide con la matriz A.

Para cada $A \in \mathbb{C}^{m\times n}$ definimos, como en el caso de las transformaciones,

$$\begin{aligned} R(A) &= \{\mathbf{y} \in \mathbb{C}^m : \mathbf{y} = A\mathbf{x} \text{ para algún } x \in \mathbb{C}^n\}, \text{ la imagen de } A, \\ N(A) &= \{\mathbf{x} \in \mathbb{C}^n : A\mathbf{x} = \mathbf{0}\}, \text{ el espacio nulo de } A. \end{aligned}$$

Análogo también a las transformaciones lineales, si A es una matriz cuadrada, diremos que es *invertible* o *no singular* si existe una matriz, que denotaremos por A^{-1}, tal que

$$A^{-1}A = I \quad \text{y} \quad AA^{-1} = I. \tag{1.3.5}$$

Para denotar una matriz $A \in \mathbb{C}^{m\times n}$ y tal que $rango(A) = r$ usaremos $A \in \mathbb{C}_r^{m\times n}$.

Proposición 1.2. *Sea $A \in \mathbb{C}^{m\times n}$ y sea $B \in \mathbb{C}^{n\times p}$ tal que $R(A) = R(AB)$, entonces existe una matriz $U \in \mathbb{C}^{p\times n}$ tal que $ABU = A$.*

Demostración. Dado $\mathbf{e}_i \in \mathbb{C}^n$, el i-ésimo vector estándar, se tiene que $A\mathbf{e}_i = \mathbf{a}_i$ que es la i-ésima columna de la matriz A. Por otro lado, por hipótesis, existe $\mathbf{u}_i \in \mathbb{C}^p$ tal que $AB\mathbf{u}_i = \mathbf{a}_i$. Sea $U = \begin{bmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n\end{bmatrix}$, luego $A = ABU$ y obtenemos el resultado deseado. □

Sea $\langle \cdot, \cdot \rangle$ el producto interno éstandar. Si $A \in \mathbb{C}^{m\times n}$ entonces

$$\langle A\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, A^*\mathbf{y}\rangle, \text{ para cada } \mathbf{x} \in \mathbb{C}^n, \mathbf{y} \in \mathbb{C}^m. \tag{1.3.6}$$

$H \in \mathbb{C}^{n\times n}$ es hermitiana si y sólo si

$$\langle H\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, H\mathbf{y}\rangle, \text{ para cada } \mathbf{x}, \mathbf{y} \in \mathbb{C}^n. \tag{1.3.7}$$

Si $\langle A\mathbf{x}, \mathbf{x}\rangle = \langle \mathbf{x}, A\mathbf{x}\rangle$ para cada $\mathbf{x}$ entonces A no necesariamente es hermitiana. Ejemplo: $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$.

Sean $\langle \cdot, \cdot \rangle_{\mathbb{C}^m}$ y $\langle \cdot, \cdot \rangle_{\mathbb{C}^n}$ productos internos en $\mathbb{C}^m$ y $\mathbb{C}^n$, respectivamente, y sea $A \in \mathcal{L}(\mathbb{C}^n, \mathbb{C}^m)$. El *adjunto* de A, denotado como A^*, es la transformación lineal $A^* \in \mathcal{L}(\mathbb{C}^m, \mathbb{C}^n)$ tal que

$$\langle A\mathbf{v}, \mathbf{u}\rangle_{\mathbb{C}^m} = \langle \mathbf{v}, A^*\mathbf{u}\rangle_{\mathbb{C}^n} \tag{1.3.8}$$

para cada $\mathbf{v} \in \mathbb{C}^n, \mathbf{u} \in \mathbb{C}^m$. A menos que se especifique lo contrario, usaremos el producto interno éstandar, en cuyo caso la *adjunta* coincide con la *conjugada transpuesta*.

Dado un subespacio L de $\mathbb{C}^n$, definimos

$$L^\perp := \{\mathbf{x} \in \mathbb{C}^n : \langle \mathbf{x}, \mathbf{y}\rangle = 0 \quad \forall \mathbf{y} \in L\}. \tag{1.3.9}$$

Es decir, el conjunto de los $\mathbf{x} \in \mathbb{C}^n$ tales que son ortogonales a cada vector en L. $L^\perp$ es un subespacio complementario de L y es llamado el *complemento ortogonal* de L. Si $M \subseteq L^\perp$ es un subespacio, entonces $L \oplus M$ es llamada la *suma directa ortogonal* de L y M y la denotaremos con $L \oplus^{\perp} M$. En particular, $\mathbb{C}^n$ es la suma directa ortogonal de L y $L^{\perp}$,

$$\mathbb{C}^n = L \oplus^{\perp} L^{\perp}. \tag{1.3.10}$$

Para cualquier matriz $A \in \mathbb{C}^{m\times n}$ podemos tomar los subespacios asociados

$$\begin{aligned} N(A), R(A^*) &\subseteq \mathbb{C}^n, \\ N(A^*), R(A) &\subseteq \mathbb{C}^m. \end{aligned}$$

Un importante resultado es que estos pares son complementos ortogonales.

Teorema 1.3. *Para cualquier matriz $A \in \mathbb{C}^{m\times n}$ se cumplen las siguientes igualdades:*

$$N(A) = R(A^*)^{\perp}, \tag{1.3.11}$$

$$N(A^*) = R(A)^{\perp}. \tag{1.3.12}$$

Demostración. Sea $\mathbf{x} \in N(A)$, entonces como consecuencia de la ecuación (1.3.8) podemos ver que $\mathbf{x} \in R(A^*)^{\perp}$. Esto prueba que $N(A) \subseteq R(A^*)^{\perp}$. Recíprocamente si $\mathbf{x} \in R(A^*)^{\perp}$ a través de la ecuación (1.3.8) podemos ver que $\langle A\mathbf{x}, \mathbf{y}\rangle = 0$ para cada $\mathbf{y} \in \mathbb{C}^m$, luego $A\mathbf{x} = \mathbf{0}$. Esto prueba que $R(A^*)^{\perp} \subseteq N(A)$. Por lo tanto $R(A^*)^{\perp} = N(A)$. Si intercambiamos los roles de A y A^* en las líneas anteriores, obtenemos $R(A)^{\perp} = N(A^*)$. □

Si $A \in \mathbb{C}^{n\times n}$, $\mathbf{x} \in \mathbb{C}^n$ distinto de cero, y $\lambda \in \mathbb{C}$ es tal que

$$A\mathbf{x} = \lambda\mathbf{x}, \tag{1.3.13}$$

entonces diremos que λ es un *valor propio* de A correspondiente al *vector propio* $\mathbf{x}$. El conjunto de valores propios de A es llamado el *espectro* de A y es denotado como $\lambda(A)$. Si λ es un valor propio de A, el subespacio

$$\{\mathbf{x} \in \mathbb{C}^n : A\mathbf{x} = \lambda\mathbf{x}\} \tag{1.3.14}$$

es llamado el *eigenespacio* (o *espacio propio*) de A correspondiente al valor propio λ, denotado por E_{λ}, y su dimensión es llamada la *multiplicidad geométrica* del valor propio λ.

Definición 1.4. *Dos matrices cuadradas A y B en $\mathbb{C}^{n\times n}$ se dirán* similares *si*

$$B = S^{-1}AS \tag{1.3.15}$$

para alguna matriz invertible S.

Definición 1.5. *Dada una matriz $A \in \mathbb{C}^{n\times n}$ definimos la traza de A como $traza(A) = a_{11} + \cdots + a_{nn}$.*

La siguiente proposición nos ayudará a ver que matrices similares coinciden en la traza.

Proposición 1.6. *Sean $A, B \in \mathbb{C}^{n\times n}$, entonces $traza(AB) = traza(BA)$.*

Demostración.

$$\begin{aligned} traza(AB) &= \sum_{i=1}^{n} AB[i,i] = \sum_{i=1}^{n}\sum_{k=1}^{n} A[i,k]B[k,i] = \sum_{k=1}^{n}\sum_{i=1}^{n} B[k,i]A[i,k] \\ &= \sum_{k=1}^{n} BA[k,k] = traza(BA) \end{aligned}$$

□

Corolario 1.7. *Si A es similar a B, entonces $traza(A) = traza(B)$.*

Demostración. Sean $A, B \in \mathbb{C}^{n\times n}$ matrices similares, luego existe una matriz $S \in \mathbb{C}^{n\times n}$ tal que $SA = BS$. Utilizando la proposición anterior obtenemos $traza(B) = traza(BSS^{-1}) = traza(S^{-1}BS) = traza(A)$. □

Definición 1.8. *Sean L y M son subespacios de $\mathbb{C}^n$ tales que $L \oplus M = \mathbb{C}^n$.*

- *Denotaremos por $P_{L,M}$ a la transformación lineal que le asigna a $\boldsymbol{x} \in \mathbb{C}^n$ su proyección en L a lo largo de M y la llamaremos* el proyector en L a lo largo de M, *o,* proyector oblicuo.

- *Si $M = L^{\perp}$, $P_{L,M}$ se denotará simplemente como P_L y se llamará* proyector ortogonal.

Observación 1.9. *Si P_L es como en la definición anterior, entonces $P_L\boldsymbol{u} = \boldsymbol{u}$ si $\boldsymbol{u} \in L$ y, $P_L\boldsymbol{u} = \boldsymbol{0}$ si $\boldsymbol{u} \in L^{\perp}$.*

Proposición 1.10. *Sean V y W espacios vectoriales, y sea $T : V \to W$ una transformación lineal. Entonces T es inyectiva si y sólo si $N(T) = \{0\}$.*

Proposición 1.11. *Sea $A \in \mathbb{F}^{m\times n}$. Si $P \in \mathbb{F}^{m\times m}$ y $Q \in \mathbb{F}^{n\times n}$ son matrices invertibles, entonces*

(a) *$rango(AQ) = rango(A)$,*

(b) *$rango(PA) = rango(A)$, y por tanto*

(c) *$rango(PAQ) = rango(A)$.*

Proposición 1.12. *Sean $T : V \to W$ y $U : W \to Z$ transformaciones lineales sobre los espacios vectoriales V, W y Z, y sean A y B matrices tales que AB está definido. Entonces*

(a) *$rango(UT) \leq rango(U)$.*

(b) *$rango(UT) \leq rango(T)$.*

(c) *$rango(AB) \leq rango(A)$.*

(d) *$rango(AB) \leq rango(B)$.*

Proposición 1.13. *Sean A y B matrices tal que AB está definida. Entonces:*

- *$R(AB) = R(A) \Leftrightarrow rango(AB) = rango(A)$,*
- *$N(AB) = N(B) \Leftrightarrow rango(AB) = rango(B)$.*

Proposición 1.14. *Si $A \in \mathbb{C}^{m\times n}$, entonces*

i) *$N(A^*A) = N(A)$.*

ii) *$R(A^*A) = R(A^*)$.*

Demostración. **i)** Basta ver que $A^*A\mathbf{x} = \mathbf{0} \Leftrightarrow A\mathbf{x} = \mathbf{0}$
$\Leftarrow$] es inmediato,
$\Rightarrow$] $A^*A\mathbf{x} = \mathbf{0} \Rightarrow A\mathbf{x} \in N(A^*) = R(A)^\perp$ por el Teorema 1.3, $\therefore A\mathbf{x} \in R(A) \cap R(A)^\perp = \{\mathbf{0}\}$.

ii) $R(A^*A) = N(A^*A)^\perp = N(A)^\perp = R(A^*)$.

□

Las demostraciones de las proposiciones anteriores no son muy complicadas y pueden encontrarse en [6].

1.4. Forma Normal de Hermite

Definición 1.15. *Una matriz en* $\mathbb{C}_r^{m\times n}$ *se dirá que está en* forma normal de Hermite *(o forma echelon reducida) si:*

- *las primeras* r *filas contienen por lo menos un elemento distinto de cero; las filas restantes contienen sólo ceros;*
- *hay* r *enteros* c_r *tales que*

$$1 \leq c_1 < c_2 < \cdots < c_r \leq n, \tag{1.4.1}$$

  y el primer elemento distinto de cero en la fila $i \in \overline{1,r}$*, aparece en la columna* c_i*; y*

- *todos los demás elementos en la columna* c_i *son cero,* $i \in \overline{1,r}$*.*

Con una permutación adecuada de sus columnas, una matriz $H \in \mathbb{C}_r^{m\times n}$ en la forma normal de Hermite puede convertirse en una matriz R con la siguiente forma

$$R = \begin{bmatrix} I_r & K \\ O & O \end{bmatrix}. \tag{1.4.2}$$

Proposición 1.16. *Sea* $A \in \mathbb{C}_r^{m\times n}$*, entonces existen* $E \in \mathbb{C}_m^{m\times m}$ *y* $P \in \mathbb{C}_n^{n\times n}$ *tales que*

$$EAP = \begin{bmatrix} I_r & K \\ O & O \end{bmatrix} \quad \text{o equivalentemente} \quad A = E^{-1} \begin{bmatrix} I_r & K \\ O & O \end{bmatrix} P^{-1}. \tag{1.4.3}$$

Hay que notar que E y P son de hecho operaciones elementales de las filas (multiplicar una fila por un escalar, sumarle α veces una fila a otra, intercambiar la fila i por la j), de la matriz A.

En [1] hay ejemplos donde se puede ver cómo transformar una matriz a su forma normal de Hermite y, una vez ahí, cómo llevarla a una matriz de la forma de (1.4.2), así como algunas propiedades extras.

Capítulo 2

Existencia y Construcción

Tomaron a uno de los que miraban, lo hicieron pedazos y en un instante lo juntaron todo y lo resucitaron. -Ea, ahora despedazaos a vosotros mismos -dijeron los Ajawab. Tomó Xbalamqué a Junajpú, lo despedazó y volvió a resucitarlo. Al ver estos prodigios los Señores pidieron ser despedazados y resucitados. Los muchachos los despedazaron y ya no volvieron a resucitarlos. Y así fueron vencidos los Ajawab de Xibalbá, los Señores del Infierno, por Junajpú e Xbalamqué.[2]

Nuestro objetivo en este capítulo es presentar la existencia y construcción de los inversos generalizados. Para ello empezaremos la siguiente sección dando tres maneras de definir la inversa generalizada (o inversa de Moore-Penrose) de una matriz y probando que son equivalentes.

2.1. La inversa de Moore-Penrose o inversa generalizada

Consideremos la siguiente ecuación:

[2]Popol Wuj, Edición de Albertina Saravia E.

$$A\mathbf{x} = \mathbf{b}, \quad A \in \mathbb{C}^{m\times n}, \mathbf{x} \in \mathbb{C}^n, \mathbf{b} \in \mathbb{C}^m. \tag{2.1.1}$$

Si $A \in \mathbb{C}^{n\times n}$ y además es invertible, entonces no hay mayor complicación para calcular (2.1.1). La única solución es $\mathbf{x} = A^{-1}\mathbf{b}$. Si A es una matriz arbitraria en $\mathbb{C}^{m\times n}$, entonces se dificulta resolver (2.1.1). Podría no tener solución, o tener una, o una infinidad. Nos gustaría poder encontrar una matriz (o matrices) C, tal que las soluciones de (2.1.1) sean de la forma $C\mathbf{b}$. Para motivar nuestra primera definición de la inversa generalizada, consideremos la siguiente ecuación:

$$y = f(x), \quad x \in \mathcal{D} \subset \mathbb{R}, \tag{2.1.2}$$

donde f es una función real con dominio $\mathcal{D}$. Suponiendo que y pertenece a la imagen de f, una manera de resolver (2.1.2) es restringir a f a un dominio más pequeño $\mathcal{D}'$ tal que $f' = f|_{\mathcal{D}'}$ sea inyectiva y después definir f'^{-1} de la imagen de f a $\mathcal{D}'$ de manera natural, es decir, $f'^{-1}(y) = x$ si $x \in \mathcal{D}'$ y $f(x) = y$. Así $f'^{-1}(y)$ es una solución de (2.1.2) para cada $y \in R(f)$.

Podemos hacer un procedimiento análogo para tratar de resolver (2.1.1). Sea $T : \mathbb{C}^n \to \mathbb{C}^m$ la transformación lineal definida por $T\mathbf{x} = A\mathbf{x}$ para $\mathbf{x} \in \mathbb{C}^n$. Notemos que si se restringimos T a $N(A)^\perp = R(A^*)$ entonces $N(T|_{R(A^*)}) = \{\mathbf{0}\}$ y así por la Proposición 1.10, $T|_{R(A^*)}$ es inyectivo, lo cual da una idea para la primera definición de inversa generalizada.

Definición 2.1. *Definición funcional de la inversa generalizada.*
Si $A \in \mathbb{C}^{m\times n}$, defínase la transformación lineal $T^\dagger : \mathbb{C}^m \to \mathbb{C}^n$ por $T^\dagger(\boldsymbol{y}) = \boldsymbol{0}$ si $\boldsymbol{y} \in R(A)^\perp$, y $T^\dagger(\boldsymbol{y}) = (T|_{R(A^*)})^{-1}\boldsymbol{y}$ si $\boldsymbol{y} \in R(A)$. La matriz de $T^\dagger$ es denotada $A^\dagger$ y llamada la inversa generalizada de A.

Notemos que $\mathbf{y} \in R(A)^\perp \Rightarrow AA^\dagger\mathbf{y} = A\mathbf{0} = \mathbf{0}$, y $\mathbf{y} \in R(A) \Rightarrow \exists\mathbf{x} \in R(A^*)$ tal que $T(\mathbf{x}) = \mathbf{y} \in R(A) \Rightarrow T^\dagger(\mathbf{y}) = \mathbf{x}$ y así $AA^\dagger\mathbf{y} = TT^\dagger(\mathbf{y}) = T(\mathbf{x}) = \mathbf{y}$. Por otra parte, si $\mathbf{x} \in R(A^*)^\perp = N(A)$ entonces $A^\dagger A\mathbf{x} = A^\dagger\mathbf{0} = \mathbf{0}$, y si $\mathbf{x} \in R(A^*) = R(A^\dagger)$ entonces $A^\dagger A\mathbf{x} = T^\dagger T(\mathbf{x}) = ((T|_{R(A^*)})^{-1}T)(\mathbf{x}) = \mathbf{x}$.

De lo anterior, y por la Observación 1.9 se desprende que $AA^\dagger$ es el proyector ortogonal de $\mathbb{C}^m$ sobre $R(A)$, mientras que $A^\dagger A$ es el proyector ortogonal de $\mathbb{C}^n$ sobre $R(A^*) = N(A)^\perp = R(A^\dagger)$.

Esto nos sugiere una segunda definición de la inversa generalizada, la cual se atribuye a E. H. Moore:

Definición 2.2. *Definición de Moore de la inversa generalizada.*
Si $A \in \mathbb{C}^{m\times n}$, entonces la inversa generalizada de A se define como la única matriz $A^\dagger$ tal que

(a) $AA^\dagger = P_{R(A)}$, *y*

(b) $A^\dagger A = P_{R(A^\dagger)}$.

La definición de Moore (dada en 1935), no fue muy conocida, al parecer debido al hecho de que no fue expresada como aquí sino al particular estilo de Moore, que usaba notación muy engorrosa (en [1] hay un apéndice donde se puede ver la notación que usaba). Penrose dio en 1955 una definición algebraica.

Definición 2.3. *Definición de Penrose de la inversa generalizada.*
Si $A \in \mathbb{C}^{m\times n}$, entonces $A^\dagger$ es la única matriz en $\mathbb{C}^{n\times m}$ tal que

(1) $AA^\dagger A = A$

(2) $A^\dagger AA^\dagger = A^\dagger$,

(3) $(AA^\dagger)^* = AA^\dagger$,

(4) $(A^\dagger A)^* = A^\dagger A$.

Las ecuaciones (1) - (4) son llamadas ecuaciones de Penrose.

Teniendo las debidas definiciones, procedamos a verificar que son equivalentes.

Teorema 2.4. *Las definiciones 2.1, 2.2 y 2.3 son equivalentes.*

Demostración. Ya probamos que si $A^\dagger$ satisface 2.1 entonces satisface las ecuaciones (a) y (b) de 2.2.
Ahora veamos que (a) y (b) implican (1) - (4) de 2.3. Por (a) $AA^\dagger = P_{R(A)} \Rightarrow AA^\dagger A = P_{R(A)}A$. Sea $\mathbf{x} \in \mathbb{C}^n$, entonces

$$A\mathbf{x} \in R(A) \Rightarrow AA^\dagger A\mathbf{x} = P_{R(A)}A\mathbf{x} = A\mathbf{x}.$$

Por lo tanto, $AA^\dagger A = A$. De manera similar, por (b) se tiene que $A^\dagger AA^\dagger = P_{R(A^\dagger)}A^\dagger = A^\dagger$. Con lo anterior, (a) y (b) implican (1) y (2). Que cumplan (3) y (4) es inmediato por ser proyectores ortogonales.

Observemos que como la definición 2.1 es por contrucción y la matriz $A^\dagger$ que construye satisface (a), (b) y (1)-(4) ya tenemos resuelto el problema de existencia en 2.2 y 2.3. A continuación probaremos que (1)-(4) implican (a) y (b). Supongamos pues, que $A^\dagger$ satisface (1)-(4). Por (2) y por (3) tenemos respectivamente

$$\begin{aligned} A^\dagger AA^\dagger = A^\dagger \quad &\Rightarrow \quad (AA^\dagger)^2 = AA^\dagger AA^\dagger = AA^\dagger, \\ (AA^\dagger)^* \quad &= \quad AA^\dagger. \end{aligned}$$

Con estas dos propiedades, $AA^\dagger$ es un proyector ortogonal. Resta ver que $R(A) = R(AA^\dagger)$. Para esto, observemos que para matrices B y C cualesquiera, $R(BC) \subseteq R(B)$, así, usando (1) tenemos que

$$R(A) = R(AA^\dagger A) \subseteq R(AA^\dagger) \subseteq R(A).$$

De esta manera, $R(AA^\dagger) = R(A)$ y por tanto $AA^\dagger = P_{R(A)}$. De manera análoga se demuestra que $A^\dagger A = P_{R(A^\dagger)}$.

Finalmente veamos que si $A^\dagger$ satisface (1)-(4), (a) y (b) entonces satisface 2.1. Con esto $A^\dagger$ será única y quedaran demostradas todas las equivalencias. En efecto, supongamos que $A^\dagger$ satisface (1)-(4), (a) y (b). Si $\mathbf{y} \in R(A)^\perp$, entonces por (a), $AA^\dagger\mathbf{y} = \mathbf{0}$. Además por (2) $A^\dagger\mathbf{y} = A^\dagger AA^\dagger\mathbf{y} = A^\dagger\mathbf{0} = \mathbf{0}$. Si $\mathbf{y} \in R(A)$, entonces existe $\mathbf{x} \in R(A^*)$ tal que $A\mathbf{x} = \mathbf{y}$, luego por (b) $A^\dagger\mathbf{y} = A^\dagger A\mathbf{x} = P_{R(A^\dagger)}(\mathbf{x})$. Además por (1)

$$AA^\dagger A = A \Rightarrow (AA^\dagger A)^* = A^* \Rightarrow (A^\dagger A)^* A^* = A^*$$

de donde por (4) se tiene que $(A^\dagger A)A^* = A^*$. Así $\mathbf{x} \in R(A^*) = R(A^\dagger AA^*) \subseteq R(A^\dagger)$. Por lo tanto, $A^\dagger\mathbf{y} = P_{R(A^\dagger)}(\mathbf{x}) = \mathbf{x}$. Pero $\mathbf{x} = (T|_{R(A^*)})^{-1}\mathbf{y}$. $\therefore$ $A^\dagger$ satisface 2.1, lo cual termina la prueba. $\square$

Debido a que tanto la definición de Moore, como la de Penrose son equivalentes, la única matriz $A^\dagger$ que las satisface es llamada también la inversa de Moore-Penrose de A. Notemos además que si A es invertible entonces

$A^{-1} = A^{\dagger}$.

En [2] podemos encontrar dos métodos para calcular explícitamente $A^{\dagger}$ (uno de los cuales se basa en la definición 2.1), así como algunos ejemplos de ello.

2.2. {1}-Inversas

Como es de esperarse, lo interesante es estudiar matrices que cumplan con una o varias de las cuatro ecuaciones de Penrose. Por ello es conveniente introducir la siguiente notación.

Definición 2.5. *Para $A \in \mathbb{C}^{m\times n}$, $A\{i, j, \dots, k\}$ denotará al conjunto de matrices $X \in \mathbb{C}^{n\times m}$ que satisfacen las ecuaciones (i), (j), ... (k) de las ecuaciones (1) - (4) de Penrose. Una matriz $X \in A\{i, j, \dots, k\}$ es llamada una $\{i, j, \dots, k\} -$ inversa de A (y denotada por $A^{(i,j,\dots,k)}$).*

Dada una matriz $R \in \mathbb{C}_r^{m\times n}$ como en (1.4.2) en la página 12, es decir

$$R = \begin{bmatrix} I_r & K \\ O & O \end{bmatrix},$$

es fácil construir una {1}-inversa: para cualquier $L \in \mathbb{C}^{(n-r)\times(m-r)}$, la matriz

$$S = \begin{bmatrix} I_r & O \\ O & L \end{bmatrix}$$

de $n \times m$, es una {1}-inversa de R ya que

$$RSR = \begin{bmatrix} I_r & KL \\ O & O \end{bmatrix} R = \begin{bmatrix} I_r & KL \\ O & O \end{bmatrix} \begin{bmatrix} I_r & K \\ O & O \end{bmatrix} = R.$$

Notemos que si R tiene rango completo de columna (r = n) entonces

$$R = \begin{bmatrix} I_r \\ O \end{bmatrix} \quad \text{y} \quad S = \begin{bmatrix} I_r & O \end{bmatrix},$$

por otra parte si R tiene rango completo de fila (r = m) entonces

$$R = \begin{bmatrix} I_r & K \end{bmatrix} \quad \text{y} \quad S = \begin{bmatrix} I_r \\ O \end{bmatrix}.$$

De esta manera, por la Proposición 1.16 en la página 12, el siguiente teorema muestra que la contrucción de $\{1\}$-inversas para una matriz arbitraria $A \in \mathbb{C}^{m\times n}$ se simplifica al modificar A para obtener una matriz en la forma normal de Hermite.

Teorema 2.6. *Sea $A \in \mathbb{C}_r^{m\times n}$, y sean $E \in \mathbb{C}_m^{m\times m}$, $P \in \mathbb{C}_n^{n\times n}$ tales que*

$$EAP \quad = \quad \begin{bmatrix} I_r & K \\ O & O \end{bmatrix}.$$

Entonces para cualquier $L \in \mathbb{C}^{(n-r)\times(m-r)}$ la matriz de $n \times m$

$$X \quad = \quad P \begin{bmatrix} I_r & O \\ O & L \end{bmatrix} E, \tag{2.2.1}$$

es una $\{1\}$-inversa de A. Las matrices EAP y X pueden ser reinterpretadas si $r = n$ o si $r = m$.

Demostración. Por (1.4.3) en la página 12 podemos expresar a A como

$$A \quad = \quad E^{-1} \begin{bmatrix} I_r & K \\ O & O \end{bmatrix} P^{-1},$$

y de esta manera

$$AXA \,=\, E^{-1} \begin{bmatrix} I_r & K \\ O & O \end{bmatrix} P^{-1}\, P \begin{bmatrix} I_r & O \\ O & L \end{bmatrix} E\, E^{-1} \begin{bmatrix} I_r & K \\ O & O \end{bmatrix} P^{-1} \,=$$

$$E^{-1} \begin{bmatrix} I_r & K \\ O & O \end{bmatrix} \begin{bmatrix} I_r & O \\ O & L \end{bmatrix} \begin{bmatrix} I_r & K \\ O & O \end{bmatrix} P^{-1} \,=\, E^{-1} \begin{bmatrix} I_r & K \\ O & O \end{bmatrix} P^{-1} \,=\, A.$$

Por lo tanto, cualquier matriz X que sea como en (2.2.1), satisface $AXA = A$, es decir, es $\{1\}$-inversa de A. $\square$

Cabe mencionar que como P y E son invertibles, por la Proposición 1.11 en la página 11:

$$rango(X) \quad = \quad rango \begin{bmatrix} I_r & O \\ O & L \end{bmatrix} \quad = \quad r \,+\, rango(L). \tag{2.2.2}$$

Como L es una matriz arbitraria, tenemos que existen $\{1\}$-inversas de A cuyo *rango* es mayor o igual que r y menor o igual que $min\{n, m\}$.

A continuación daremos un lema que nos muestra ciertas propiedades de las {1}-inversas. Para esto, hagamos las siguientes aclaraciones:

- Dada una matriz A, $A^{(1)}$ denotará una de sus {1}-inversas cualquiera.
- Para cualquier escalar λ definamos

$$\lambda^{\dagger} = \begin{cases} \lambda^{-1}, & \text{si } \lambda \neq 0; \\ 0, & \text{si } \lambda = 0. \end{cases} \tag{2.2.3}$$

- Una matriz cuadrada E se dirá idempotente si $E^2 = E$.

Lema 2.7. *Sean $A \in \mathbb{C}_r^{m\times n}$ y $\lambda \in \mathbb{C}$. Entonces:*

(a) $(A^{(1)})^* \in A^*\{1\}$.

(b) *Si A es invertible,* $A^{(1)} = A^{-1}$.

(c) $\lambda^{\dagger} A^{(1)} \in (\lambda A)\{1\}$.

(d) $rango(A^{(1)}) \geq rango(A)$.

(e) *Si S y T son invertibles,* $T^{-1}A^{(1)}S^{-1} \in SAT\{1\}$.

(f) *$AA^{(1)}$ y $A^{(1)}A$ son idempotentes y tienen el mismo rango que A.*

Demostración. (a) Basta ver que $A^*(A^{(1)})^*A^* = A^*$. En efecto,

$$AA^{(1)}A = A \Rightarrow A^* = (AA^{(1)}A)^* = A^*(AA^{(1)})^* = A^*(A^{(1)})^*A^*.$$

(b) Supongamos que A es invertible, así existe A^{-1}, como $AA^{(1)}A = A$ entonces

$$A^{-1} = A^{-1}AA^{-1} = A^{-1}(AA^{(1)}A)A^{-1} = (A^{-1}A)A^{(1)}(AA^{-1}) = A^{(1)}.$$

Por lo tanto $A^{-1} = A^{(1)}$.
(c) $\lambda^{\dagger}A^{(1)} \in (\lambda A)\{1\}$ ya que

$$\lambda A(\lambda^{\dagger}A^{(1)})\lambda A = \lambda A(\lambda^{\dagger}\lambda)A^{(1)}A = \lambda AA^{(1)}A = \lambda A.$$

(d) Por la Proposición 1.12 en la página 11

$$rango(A) = rango(AA^{(1)}A) \leq rango(A^{(1)}A) \leq rango(A^{(1)}).$$

(e) Si S y T son invertibles $T^{-1}A^{(1)}S^{-1} \in SAT\{1\}$ ya que

$$SAT(T^{-1}A^{(1)}S^{-1})SAT = SA(TT^{-1})A^{(1)}(S^{-1}S)AT = SAA^{(1)}AT = SAT.$$

(f) $AA^{(1)}$ y $A^{(1)}A$ son idempotentes ya que $(AA^{(1)})^2 = AA^{(1)}AA^{(1)} = AA^{(1)}$ y $(A^{(1)}A)^2 = A^{(1)}AA^{(1)}A = A^{(1)}A$. Además, nuevamente por la Proposición 1.12:

$$rango(A) = rango(AA^{(1)}A) \leq rango(AA^{(1)}) \leq rango(A),$$

$$rango(A) = rango(AA^{(1)}A) \leq rango(A^{(1)}A) \leq rango(A).$$

Por lo tanto $rango(AA^{(1)}) = rango(A) = rango(A^{(1)}A)$.

□

Corolario 2.8. *Si $A^{(1)} \in A\{1\}$, entonces*

- $R(AA^{(1)}) = R(A)$,
- $N(AA^{(1)}) = N(A)$,
- $R((AA^{(1)})^*) = R(A^*)$.

Demostración. Son consecuencia inmediata de (a) y (f) del lema anterior y de la Proposición 1.13. □

El siguiente lema nos dice que si $A \in \mathbb{C}_r^{m\times n}$ tiene rango completo de columna, entonces sus $\{1\}$-inversas son sus inversas por la izquierda. Si A tiene rango completo de fila, sus $\{1\}$-inversas son sus inversas por la derecha.

Lema 2.9. *Sean $A \in \mathbb{C}_r^{m\times n}$. Entonces:*

(a) $A^{(1)}A = I_n$ *si y sólo si* $r = n$.

(b) $AA^{(1)} = I_m$ *si y sólo si* $r = m$.

Demostración. (b) $\Leftarrow$] Supongamos que $r = m$. Notemos que $AA^{(1)} \in \mathbb{C}^{m\times m}$, además por (f) del lema anterior $rango(A^{(1)}A) = rango(A) = m$ y $(AA^{(1)})^2 = AA^{(1)}$. Entonces $AA^{(1)}$ es invertible y

$$(AA^{(1)})^2(AA^{(1)})^{-1} = AA^{(1)}(AA^{(1)})^{-1}$$

de donde $AA^{(1)} = I_m$.
$\Rightarrow$] $AA^{(1)} = I_m \Rightarrow rango(AA^{(1)}) = I_m = m$. Entonces por (f) del lema anterior $rango(A) = m$.
(a) Se prueba de manera similar. □

2.3. {1,2}-Inversas

Bjerhammar fue el primero en observar que la existencia de una {1}-inversa de una matriz A implica la existencia de una {1,2}-inversa.

Lema 2.10. *Sean $Y, Z \in A\{1\}$, y sea*

$$X = YAZ.$$

Entonces $X \in A\{1,2\}$.

Demostración.

$$AXA = A(YAZ)A = (AYA)ZA = AZA = A,$$

$$XAX = (YAZ)A(YAZ) = Y(AZA)YAZ = YAYAZ = YAZ = X.$$

$\square$

Por la simetría de las ecuaciones (1) y (2), $X \in A\{1,2\}$ si y sólo si $A \in X\{1,2\}$. Cuando esto ocurra, diremos que A y X son {1,2}-inversas la una de la otra.

Observemos también que si A y X son {1,2}-inversas la una de la otra, entonces por (f) del lema 2.7

$$rango(A) = rango(AXA) \leq rango(AX) \leq rango(A)$$

$$rango(X) = rango(XAX) \leq rango(AX) \leq rango(X)$$

lo cual nos dice que A y X tienen el mismo *rango*. Bjerhammar también notó que si X es una {1}-inversa de A tal que tiene el mismo rango que A, entonces X es una {1,2}-inversa de A.

Teorema 2.11. (Bjerhammar).
Dados A y $X \in A\{1\}$. $X \in A\{1,2\}$ si y sólo si $rango(X) = rango(A)$.

Demostración. $\Leftarrow$] Supogamos que $rango(X) = rango(A)$. Por (f) del lema 2.7, $rango(XA) = rango(A) = rango(X)$. Entonces $R(XA) = R(X)$ por la Proposición 1.13. Luego, por la Proposición (1.2)

$$XAY = X$$

para alguna Y. Multiplicando por A tenemos

$$AX = AXAY = AY$$

y luego multiplicando por X tenemos

$$XAX = XAY = X.$$

$\Rightarrow$] Se sigue de lo observado anteriormente. $\square$

Corolario 2.12. *Cualesquiera dos de las tres ecuaciones siguientes implica la tercera:*

$$X \in A\{1\}$$

$$X \in A\{2\}$$

$$rango(X) = rango(A).$$

Demostración. Por el teorema anterior sólo resta ver que $X \in A\{2\}$ y $rango(X) = rango(A) \Rightarrow X \in A\{1\}$. Como $X \in A\{2\}$ entonces $XAX = X$, así $A \in X\{1\}$. Entonces por el teorema anterior $A \in X\{1,2\}$ y por lo tanto A y X son {1,2}-inversas la una de la otra. $\square$

Por el Teorema 2.11, (2.2.2) muestra que la {1}-inversa que se obtiene de la forma normal de Hermite es una {1,2}-inversa si hacemos $L = O$. En otras palabras,

$$X = P \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} E \tag{2.3.1}$$

es una {1,2}-inversa de A donde P y E son invertibles y satisfacen (1.4.3) de la página 12.

2.4. {1,2,3}-, {1,2,4}-, y {1,2,3,4}-inversas.

Así como Bjerhammar probó que la existencia de una {1}-inversa implica la existencia de una {1,2}-inversa, Urquhart mostró que la existencia de una {1}-inversa de una matriz finita con entradas en $\mathbb{C}$ implica la existencia de una {1,2,3}-inversa y una {1,2,4}-inversa de dicha matriz. Sin embargo, para mostrar que $A\{1,2,3\}$ y $A\{1,2,4\}$ son no vacíos para una matriz A dada, no usaremos la {1}-inversa de A, sino la de una matriz relacionada.

Teorema 2.13. (Urquhart).
Para cualquier matriz finita A con entradas complejas, se tiene

$$Y = (A^*A)^{(1)}A^* \in A\{1,2,3\} \quad y \tag{2.4.1}$$

$$Z = A^*(AA^*)^{(1)} \in A\{1,2,4\}. \tag{2.4.2}$$

Demostración. Por la Proposición 1.14 en la página 11, $R(A^*A) = R(A^*)$, entonces $A^* = A^*AU$ para algún U, luego

$$A = U^*A^*A \Rightarrow AYA = U^*A^*A(A^*A)^{(1)}A^*A = U^*A^*A = A$$

por lo tanto $Y \in A\{1\}$. Ahora podemos aplicar a Y (d) del Lema 2.7 y obtenemos $rango(Y) \geq rango(A)$ y ahora por como Y está definido se tiene $rango(Y) \leq rango(A^*) = rango(A)$. Por lo tanto $rango(Y) = rango(A)$, y así por el Teorema 2.11 $Y \in A\{1,2\}$. Finalmente

$$AY = U^*A^*A(A^*A)^{(1)}A^* = U^*A^*A(A^*A)^{(1)}A^*AU = U^*A^*AU \Rightarrow$$

$$(AY)^* = (U^*A^*AU)^* = (AU)^*(U^*A^*)^* = U^*A^*AU = AY$$

por lo tanto $Y \in A\{1,2,3\}$. De forma similar $Z \in A\{1,2,4\}$. □

El siguiente teorema nos dice que teniendo una {1,4}-inversa y una {1,3}-inversa de una matriz finita con entradas complejas podemos conseguir la inversa de Moore-Penrose.

Teorema 2.14. (Urquhart).
Para cualquier matriz finita con entradas complejas A. se tiene que

$$A^{(1,4)}AA^{(1,3)} = A^\dagger. \tag{2.4.3}$$

Demostración. Definamos $X = A^{(1,4)}AA^{(1,3)}$, por el Lema 2.10 tenemos que $X \in A\{1,2\}$ además

$$AX = AA^{(1,4)}AA^{(1,3)} = AA^{(1,3)} \quad \text{y}$$

$$XA = A^{(1,4)}AA^{(1,3)}A = A^{(1,4)}A,$$

entonces

$$(AX)^* = (AA^{(1,3)})^* = AA^{(1,3)} \quad \text{y} \quad (XA)^* = (A^{(1,4)}A)^* = A^{(1,4)}A.$$

Por lo tanto $X \in A\{1,2,3,4\}$. Y como ya vimos que $A^\dagger$ es única, entonces $X = A^\dagger$. □

2.5. Fórmula Explícita Para $A^\dagger$

Aparentemente (en 1959), C. C. MacDuffee fue el primero en observar que la factorización de rango completo de una matriz A conduce a una fórmula explícita de su inversa de Moore-Penrose (ver [1]).

Teorema 2.15. (MacDuffee).
Si $A \in \mathbb{C}_r^{m\times n}$, $r > 0$ tiene como factorización de rango completo

$$A = FG,$$

entonces

$$A^\dagger = G^*(F^*AG^*)^{-1}F^*.$$

Demostración. Primero demostraremos que F^*AG^* es invertible. Como

$$A = FG \Rightarrow F^*AG^* = (F^*F)(GG^*)$$

que es un producto de matrices de $r \times r$, además por la Proposición 1.14 en la página 11,

$$rango(F^*F) = r = rango(GG^*)$$

por lo tanto F^*AG^* es invertible. Ahora $(F^*AG^*)^{-1} = (GG^*)^{-1}(F^*F)^{-1}$, si tomamos a

$$X = G^*(GG^*)^{-1}(F^*F)^{-1}F^*$$

se puede verificar que $X \in A\{1, 2, 3, 4\}$. □

Proposición 2.16. *Muestre que si H es hermítica e idempotente entonces $H = H^\dagger$.*

Demostración. Como H es Hermítica e idempotente se tiene que $H^* = H$ y $HH = H$. Así $(HH)^* = H^*H^* = HH$ por un lado, y por otro $HHH = (HH)H = HH = H$. Por lo tanto $H \in H\{1, 2, 3, 4\}$ y así $H = H^\dagger$. □

En [1] se pueden encontrar propiedades adicionales que cumplen las inversas generalizadas que hemos estudiado en este capítulo, así como la construcción de $\{2\}$-inversas, de una matriz $A \in \mathbb{C}_r^{m\times n}$, con *rango* entre 0 y r.

Capítulo 3

Sistemas Lineales y Caracterizaciones

Agucé la razón tanto, que oscura fue para los demás mi vida, mi pasión y mi locura.[3]

3.1. Solución de Sistemas Lineales

La principal aplicación de las $\{1\}$-inversas es en la solución de sistemas lineales, donde son usadas casi de manera análoga a como se usan las inversas usuales (cuando éstas existen). El siguiente teorema y su demostración se los debemos a Penrose.

Teorema 3.1. (Penrose).
Sean $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{p\times q}$, $D \in \mathbb{C}^{m\times q}$. Entonces la ecuación matricial

$$AXB = D \tag{3.1.1}$$

tiene solución si y sólo si para algún $A^{(1)}$ y algún $B^{(1)}$ se tiene

$$AA^{(1)}DB^{(1)}B = D, \tag{3.1.2}$$

[3]Epitafio a Jorge Cuesta por Xavier Villaurrutia.

en cuyo caso la solución general es

$$X = A^{(1)}DB^{(1)} + Y - A^{(1)}AYBB^{(1)} \tag{3.1.3}$$

para cualquier $Y \in \mathbb{C}^{n\times p}$.

Demostración. $\Leftarrow$] Supongamos que $AA^{(1)}DB^{(1)}B = D$, entonces $X = A^{(1)}DB^{(1)}$ es una solución de (3.1.1).
$\Rightarrow$] Supongamos que X es una solución de (3.1.1), entonces

$$D = AXB = AA^{(1)}AXBB^{(1)}B = AA^{(1)}DB^{(1)}B.$$

Ahora veamos que si $Y \in \mathbb{C}^{n\times p}$ y se satisface (3.1.2), entonces $X = A^{(1)}DB^{(1)} + Y - A^{(1)}AYBB^{(1)}$ es solución de (3.1.1). En efecto,

$$AXB = AA^{(1)}DB^{(1)}B + AYB - AA^{(1)}AYBB^{(1)}B = D + AYB - AYB = D.$$

Por otra parte, si X es una solución cualquiera de (3.1.1), entonces claramente

$$X = A^{(1)}DB^{(1)} + X - A^{(1)}AXBB^{(1)},$$

la cual es de la forma (3.1.3). $\square$

La siguiente caracterización del conjunto $A\{1\}$, en términos de $A^{(1)}$, un elemento arbitrario del conjunto, es debido a Bjerhammar.

Corolario 3.2. *Sean* $A \in \mathbb{C}^{m\times n}$, $A^{(1)} \in A\{1\}$. *Entonces*

$$A\{1\} = \{A^{(1)} + Z - A^{(1)}AZAA^{(1)} : Z \in \mathbb{C}^{n\times m}\}. \tag{3.1.4}$$

Demostración. Sabemos que $A\{1\} = \{X : AXA = A\}$ es decir, $A\{1\}$ es el conjunto solución de la ecuación $AXA = A$. Por lo tanto, si en el teorema anterior hacemos $Y = A^{(1)} + Z$ tenemos que la solución general de la ecuación es

$$\begin{aligned} X &= A^{(1)}AA^{(1)} + A^{(1)} + Z - [A^{(1)}AA^{(1)}AA^{(1)} + A^{(1)}AZAA^{(1)}] \\ &= A^{(1)}AA^{(1)} + A^{(1)} + Z - A^{(1)}AA^{(1)} - A^{(1)}AZAA^{(1)} \\ &= A^{(1)} + Z - A^{(1)}AZAA^{(1)} \end{aligned}$$

por lo tanto $A\{1\} = \{A^{(1)} + Z - A^{(1)}AZAA^{(1)} : Z \in \mathbb{C}^{n\times m}\}$. $\square$

El siguiente corolario, resulta del Teorema 3.1, aplicado a sistemas ordinarios de ecuaciones lineales.

Corolario 3.3. *Sean $A \in \mathbb{C}^{m\times n}$, $\boldsymbol{b} \in \mathbb{C}^m$. Entonces la ecuación*

$$A\boldsymbol{x} = \boldsymbol{b} \tag{3.1.5}$$

tiene solución si y sólo si, para algún $A^{(1)}$,

$$AA^{(1)}\boldsymbol{b} = \boldsymbol{b}, \tag{3.1.6}$$

en cuyo caso la solución general de (3.1.5) *es*

$$X = A^{(1)}b + (I - A^{(1)}A)\boldsymbol{y} \tag{3.1.7}$$

para $\boldsymbol{y} \in \mathbb{C}^n$ arbitraria.

Demostración. Tomando en el Teorema 3.1 a $B = I_1$ y a $D = b$, se obtiene el resultado. □

Ahora presentaremos una caracterización alternativa de $A\{1\}$. Esta caracterización aparece en la tesis doctoral de C. A. Rohde, el cual atribuye el resultado a R. C. Bose (ver [1]).

Teorema 3.4. *Sean $A \in \mathbb{C}^{m\times n}$, $X \in \mathbb{C}^{n\times m}$. Entonces $X \in A\{1\}$ si y sólo si para cualquier $\boldsymbol{b}$ tal que $A\boldsymbol{x} = \boldsymbol{b}$ admita solución, se tiene que $\boldsymbol{x} = X\boldsymbol{b}$ es una solución.*

Demostración. $\Leftarrow$] Sea $\mathbf{a}_j$ la j-ésima columna de la matriz A, entonces el sistema $A\mathbf{x} = \mathbf{a}_j$ tiene solución ya que $A\mathbf{e}_j = a_j$, donde $\mathbf{e}_j$ es el j-ésimo vector unitario. Entonces $X\mathbf{a}_j$ es una solución, es decir $AX\mathbf{a}_j = \mathbf{a}_j$ $(j \in \overline{1,n})$. Por lo tanto $AXA = A$.
$\Rightarrow$] Se sigue inmediatamente por el corolario anterior y (3.1.6). □

La siguiente proposición nos dice cuándo dos ecuaciones matriciales tienen una solución en común.

Proposición 3.5. *Las ecuaciones matriciales*

$$AX = B, \quad XD = E \tag{3.1.8}$$

tienen una solución común si y sólo si cada ecuación tiene una solución por su cuenta y

$$AE = BD.$$

Demostración. $\Leftarrow$] Supongamos que $AX = B$ tiene solución, $X'D = E$ tiene solución, y $AE = BD$. Entonces por el Teorema 3.1,

$$AA^{(1)}B = B \quad \text{y} \quad ED^{(1)}D = E.$$

Afirmamos que $A^{(1)}B + ED^{(1)} - A^{(1)}AED^{(1)}$ es solución común de las ecuaciones (3.1.8). En efecto,

$$AX = A(A^{(1)}B+ED^{(1)}-A^{(1)}AED^{(1)}) = AA^{(1)}B+AED^{(1)}-AA^{(1)}AED^{(1)} =$$

$$= AA^{(1)}B = B,$$

por otra parte, como $AE = BD$,

$$XD = (A^{(1)}B+ED^{(1)}-A^{(1)}BDD^{(1)})D = A^{(1)}BD+ED^{(1)}D-A^{(1)}BDD^{(1)}D =$$

$$= ED^{(1)}D = E.$$

Por lo tanto X es solución común de (3.1.8).
$\Rightarrow$] Sea X la solución común de $AX = B$ y $XD = E$, entonces

$$XD = E \Rightarrow AXD = AE \quad \text{y}$$

$$AX = B \Rightarrow AXD = BD.$$

Por lo tanto, $AE = BD$. □

3.2. Caracterización de $A\{1,3\}$ y $A\{1,4\}$.

Ahora nuestro interés se centra en caracterizar al conjunto $A\{1,3\}$. El siguiente teorema nos ayudará a alcanzar ese fin.

Teorema 3.6. *El conjunto $A\{1,3\}$ consiste de todas las X que son solución de*

$$AX = AA^{(1,3)}, \tag{3.2.1}$$

donde $A^{(1,3)}$ es un elemento arbitrario de $A\{1,3\}$.

Demostración. Si X satisface (3.2.1), entonces $AXA = AA^{(1,3)}A = A$, y además

$$(AX)^* = (AA^{(1,3)})^* = AA^{(1,3)} = AX.$$

Por lo tanto $X \in A\{1,3\}$. Por otra parte, usando (a) del Lema 2.7, si $X \in A\{1,3\}$ entonces

$$\begin{aligned} AA^{(1,3)} &= AXAA^{(1,3)} = (AX)^*AA^{(1,3)} = (AX)^*(AA^{(1,3)})^* \\ &= X^*A^*(A^{(1,3)})^*A^* = X^*A^* = (AX)^* = AX. \end{aligned}$$

Por lo tanto, X satisface (3.2.1). $\square$

Ahora, usando este resultado y el Teorema 3.1, obtenemos una caracterización para $A\{1,3\}$.

Corolario 3.7. *Sean $A \in \mathbb{C}^{m\times n}$, $A^{(1,3)} \in A\{1,3\}$. Entonces*

$$A\{1,3\} = \{A^{(1,3)} + (I - A^{(1,3)}A)Z : Z \in \mathbb{C}^{n\times m}\}.$$

Demostración. Por el teorema anterior $A\{1,3\} = \{X : AX = AA^{(1,3)}\}$, de esta manera, usando el Teorema 3.1 la forma general de X es

$$X = A^{(1,3)}AA^{(1,3)} + Y - A^{(1,3)}AY,$$

si ahora hacemos $Y = A^{(1,3)} + Z$, obtenemos

$$\begin{aligned} X &= A^{(1,3)}AA^{(1,3)} + (A^{(1,3)} + Z) - A^{(1,3)}A(A^{(1,3)} + Z) = \\ &= A + A^{(1,3)} + Z - A^{(1,3)}AA^{(1,3)} - A^{(1,3)}AZ = \\ &= A^{(1,3)} + Z - A^{(1,3)}AZ = A^{(1,3)} + (I - A^{(1,3)}A)Z. \end{aligned}$$

Por lo tanto $A\{1,3\} = \{A^{(1,3)} + (I - A^{(1,3)}A)Z : Z \in \mathbb{C}^{n\times m}\}$. $\square$

El siguiente teorema y su corolario, se obtienen de manera similar a las demostraciones del Teorema 3.6 y el Corolario 3.7.

Teorema 3.8. *El conjunto $A\{1,4\}$ consiste de todas las X que son solución de*

$$XA = A^{(1,4)}A.$$

Corolario 3.9. *Sean $A \in \mathbb{C}^{m\times n}$, $A^{(1,4)} \in A\{1,4\}$. Entonces*

$$A\{1,4\} = \{A^{(1,4)} + Y(I - AA^{(1,4)}) : Y \in \mathbb{C}^{n\times m}\}.$$

3.3. Caracterización de $A\{2\}$ y otros de sus subconjuntos.

Como la ecuación $XAX = X$ es no lineal, no podemos simplemente aplicar el Teorema 3.1 para encontrar una caracterización de $A\{2\}$. Sin embargo, podemos encontrar una caracterización usando una factorización de rango completo de X.

Denotaremos por $A\{i, j, \ldots, k\}_s$ al subconjunto de $A\{i, j, \ldots, k\}$ consistente de matrices de rango s.

Los conjuntos $A\{2\}_0$, $A\{2, 3\}_0$, $A\{2, 4\}_0$, $A\{2, 3, 4\}_0$, son iguales y contienen sólo un elemento. Para $A \in \mathbb{C}^{m\times n}$, este único elemento es la matriz cero de $n \times m$. Debido a esto, en el resto de esta sección supondremos que $s > 0$.

El siguiente teorema ha sido presentado por G. W. Steward. Él se lo atribuye a R. E. Funderlic (ver [1]).

Teorema 3.10. *Sean $A \in \mathbb{C}_r^{m\times n}$ y $0 < s \leq r$. Entonces*

$$A\{2\}_s = \{YZ : Y \in \mathbb{C}^{n\times s}, Z \in \mathbb{C}^{s\times m}, ZAY = I_s\}. \tag{3.3.1}$$

Demostración. Sean $Y \in \mathbb{C}^{n\times s}, Z \in \mathbb{C}^{s\times m}$ tal que $ZAY = I_s$, entonces

$$s \geq rango(Y) \geq rango(ZAY) = rango(I_s) = s$$

es decir, $rango(Y) = s$, análogamente se observa que $rango(Z) = s$. Sea $X = YZ = YI_sZ$, entonces por el Lema 2.9,

$$Y^{(1)}XZ^{(1)} = Y^{(1)}YZZ^{(1)} = I_sI_s = I_s,$$

entonces

$$\begin{aligned} rango(X) &\geq rango(Y)^{(1)}XZ^{(1)} = rango(I_s) = s = rango(I_s) \geq \\ &\geq rango(YI_sZ) = rango(X), \end{aligned}$$

por lo tanto $rango(X) = s$. Además $XAX = YZAYZ = YI_sZ = YZ = X$. Con lo cual $X \in A\{2\}_s$.
Por otra parte, consideremos $X \in A\{2\}_s$ y tomemos una factorización de

rango completo para X, es decir, $Y \in \mathbb{C}_s^{n\times s}, Z \in \mathbb{C}_s^{s\times m}$ tal que $X = YZ$, entonces $YZAYZ = YZ$. Luego, multiplicando por $Y^{(1)}$ y $Z^{(1)}$ cualesquiera, obtenemos

$$Y^{(1)}YZAYZZ^{(1)} = Y^{(1)}YZZ^{(1)}$$

y, nuevamente por el Lema 2.9 se tiene que $I_sZAYI_s = I_sI_s$, es decir, $ZAY = I_s$. □

Corolario 3.11. *Sea $A \in \mathbb{C}_r^{m\times n}$. Entonces*

$$A\{1,2\} \;=\; \{YZ : Y \in \mathbb{C}^{n\times r}, Z \in \mathbb{C}^{r\times m}, ZAY = I_r\}.$$

Demostración. por el Teorema 2.11,

$$A\{1,2\} \;=\; A\{2\}_r.$$

□

La igualdad $ZAY = I_s$ de (3.3.1), implica que $Z \in (AY)\{1,2,4\}$.. Esto nos sugiere un acercamiento para la caracterización de $A\{2,3\}$ en la cual se basa el siguiente teorema.

Teorema 3.12. *Sean $A \in \mathbb{C}_r^{m\times n}$ y $0 < s \leq r$. Entonces*

$$A\{2,3\}_s = \{Y(AY)^\dagger : AY \in \mathbb{C}_s^{m\times s}\}. \tag{3.3.2}$$

Demostración. Sea $X = Y(AY)^\dagger$, donde $AY \in \mathbb{C}_s^{m\times s}$. Entonces $AX = AY(AY)^\dagger$. Por (3) de las ecuaciones de Penrose se tiene que

$$(AX)^* = (AY(AY)^\dagger)^* = AY(AY)^\dagger = AX$$

así que $X \in A\{3\}$. Además

$$XAX = Y(AY)^\dagger AY(AY)^\dagger = Y(AY)^\dagger = X.$$

Por lo tanto $X \in A\{2,3\}$. Finalmente, como $X \in A\{2\}$ entonces $A \in X\{1\}$ y así, por el (f) del Lema 2.7

$$\begin{aligned} rango(X) &= rango(AX) = rango(AY(AY)^\dagger) = \\ &= rango(AY) = s. \end{aligned}$$

Por lo tanto $X \in A\{2,3\}_s$.
Por otra parte, sea $X \in A\{2,3\}_s$. Entonces AX es Hermítica, Idempotente

y, por (f) del Lema 2.7, tiene rango s, ya que $A \in X\{1\}$. Por la Proposición 2.16

$$(AX)^{\dagger} = AX,$$

y de esta manera

$$X(AX)^{\dagger} = XAX = X.$$

Por lo tanto $X \in \{Y(AY)^{\dagger} : AY \in \mathbb{C}_s^{m\times s}\}$. $\square$

El siguiente teorema se demuestra de manera similar.

Teorema 3.13. *Sean $A \in \mathbb{C}_r^{m\times n}$ y $0 < s \leq r$. Entonces*

$$A\{2,4\}_s = \{(YA)^{\dagger}Y : YA \in \mathbb{C}_s^{s\times m}\}.$$

3.4. Matrices Idempotentes y Proyectores

En esta sección mostraremos algunas propiedades de las matrices idempotentes y su relación con los proyectores, que como ya vimos aparecen en la definición de Moore de inversa generalizada. Para esto, comencemos con los siguientes lemas.

Lema 3.14. *Sea $E \in \mathbb{C}^{n\times n}$ idempotente. Entonces:*

(a) *E^* y $I - E$ son idempotentes.*

(b) *Los valores propios de E son 0 y 1. La multiplicidad del valor propio 1 es rango(E).*

(c) *$rango(E) = trazaE$.*

(d) *$E(I - E) = (I - E)E = O$.*

(e) *$E\boldsymbol{x} = \boldsymbol{x}$ si y sólo si $\boldsymbol{x} \in R(E)$.*

(f) *$E \in E\{1,2\}$.*

(g) *$N(E) = R(I - E)$.*

Demostración. (a)
$$E^*E^* = (EE)^* = E^*,$$
$$(I-E)(I-E) = I(I-E) - E(I-E) = I - E - E + EE = I - E.$$
(b) Sea $\mathbf{x} \in \mathbb{C}^n - \{\mathbf{0}\}$ tal que $E\mathbf{x} = \lambda\mathbf{x}$, entonces

$$\lambda\mathbf{x} = E\mathbf{x} = EE\mathbf{x} = E(\lambda\mathbf{x}) = \lambda E\mathbf{x} = \lambda^2\mathbf{x} \Leftrightarrow \lambda^2 = \lambda,$$

λ es de la forma $\lambda = a+ib$ tal que $a, b \in \mathbb{R}$, entonces $(a+ib)^2 = a^2+2iab-b^2 = a+ib$ si y sólo si $a^2 - b^2 = a \quad \text{y} \quad 2ab = b$. Si $b \neq 0$ entonces $a = 1/2$ y por tanto $1/4 - b^2 = 1/2$, es decir, $b^2 = -1/4$. Lo cual es una contracción. De esta manera, $b = 0$ y por lo tanto $\lambda = a = a^2 = \lambda^2$ de donde $\lambda = 0$ o $\lambda = 1$. Para ver que la multiplicidad del valor propio 1 es $rango(E)$, supongamos que $E \neq O$ (Si $E = O$ entonces $1 \notin \lambda(E)$). Así, existe $\mathbf{y}$ tal que $E\mathbf{y} \neq \mathbf{0}$, esto implica que $E(E\mathbf{y}) = E\mathbf{y}$ y por lo tanto $1 \in \lambda(E)$. Si verificamos que $E_1 = R(E)$, se tiene que

$$mult(1) = dim(E_1) = dim(R(E)) = rango(E).$$

En efecto, sea $\mathbf{x} \in E_1$ entonces $E\mathbf{x} = \mathbf{x} \in R(E)$. Por otra parte si $\mathbf{y} \in R(E)$ entonces existe $\mathbf{x}$ tal que $E\mathbf{x} = \mathbf{y}$ y de aquí que $E\mathbf{x} = E(E\mathbf{x}) = E\mathbf{y} = \mathbf{y}$, por lo tanto $\mathbf{y} \in E_1$.
(c) Supongamos que $N(E) = \{\mathbf{0}\}$ entonces por (1.3.1) y por (b) obtenemos $dim(\mathbb{C}^n) = dim(R(E)) = dim(E_1)$. Así $\mathbb{C}^n$ tiene una base de vectores propios, asociados al valor propio 1, de donde obtenemos que E es similar a la matriz identidad y en consecuencia $traza(E) = traza(I) = n = dim(R(E)) = rango(E)$.
Supogamos ahora que $N(E) \neq \{\mathbf{0}\}$, entonces $0 \in \lambda(E)$ y $E_0 = N(E)$, luego por (1.3.1) y por (b) se tiene $dim(E_1) + dim(E_0) = dim(\mathbb{C}^n) = n$. Sea $\mathbf{x} \in E_1 \cap E_0$, luego $E\mathbf{x} = \mathbf{x}$ y $E\mathbf{x} = \mathbf{0}$, así $x = \mathbf{0}$. Por lo tanto $E_1 \cap E_0 = \{\mathbf{0}\}$. Sea $r = dim(E_1)$ y sean $\beta_1 = \{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$, $\beta_2 = \{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ bases de E_1 y E_0 respectivamente. Sean $a_1, \ldots, a_n$ escalares tales que

$$a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \mathbf{0},$$

luego
$$a_1\mathbf{v}_1 + \cdots + a_r\mathbf{v}_r = -a_{r+1}\mathbf{v}_{r+1} - \cdots - a_n\mathbf{v}_n \in E_1 \cap E_0,$$

y así
$$a_1\mathbf{v}_1 + \cdots + a_r\mathbf{v}_r = \mathbf{0},$$

de donde obtenemos $a_1 = \cdots = a_r = 0$, dado que β_1 es linealmente independiente. Análogamente obtenemos $a_{r+1} = \cdots = a_n = 0$, por lo tanto $\beta_1 \cup \beta_0$ es linealmente independiente y consta de n vectores propios de E, luego E es similar a la matriz
$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$
Así $traza(E) = r = dim(E_1) = dim(R(E)) = rango(E)$.
(d)
$$E(I-E)\mathbf{x} = E(\mathbf{x} - E\mathbf{x}) = E\mathbf{x} - E\mathbf{x} = \mathbf{0},$$
$$(I-E)E\mathbf{x} = (E\mathbf{x} - EE\mathbf{x}) = E\mathbf{x} - E\mathbf{x} = \mathbf{0}.$$
(e) En (b) observamos que $E_1 = R(E)$, con esto la bicondicional se observa de manera inmediata.
(f) Es inmediato ($EEE = EE = E$).
(g) Sea $\mathbf{x} \in R(I-E)$ entonces existe $\mathbf{y}$ tal que $\mathbf{x} = (I-E)\mathbf{y} = \mathbf{y} - E\mathbf{y}$, entonces $E\mathbf{x} = E\mathbf{y} - EE\mathbf{y} = \mathbf{0}$. De donde $\mathbf{x} \in N(E)$ y por tanto $R(I-E) \subset N(E)$. Por otra parte, si $\mathbf{x} \in N(E)$ entonces por el Corolario 3.3 como $E\mathbf{x} = \mathbf{0}$ tiene solución, entonces para algún $E^{(1)}$, $EE^{(1)}\mathbf{0} = \mathbf{0}$ (pero por (f), $E \in E\{1\}$ y E cumple que $EE\mathbf{0} = \mathbf{0}$), y $\mathbf{x}$ es de la forma $\mathbf{x} = E^{(1)}\mathbf{0} + (I - E^{(1)}E)\mathbf{y}$, para $\mathbf{y}$ arbitrario. Entonces, $\mathbf{x} = E\mathbf{0} + (I - EE)\mathbf{y} = (I-E)\mathbf{y}$, para $\mathbf{y}$ arbitrario. Por lo tanto $\mathbf{x} \in R(I-E)$ y así $N(E) \subset R(I-E)$. □

Lema 3.15. *Sea una matriz cuadrada con factorización de rango completo*
$$E = FG.$$
Entonces E es idempotente si y sólo si $GF = I$.

Demostración. Si $GF = I$, entonces claramente
$$(FG)^2 = FGFG = FIG = FG$$
Por otra parte, como F tiene rango completo de columna y G tiene rango completo de fila, por el Lema 2.9 tenemos
$$F^{(1)}F = GG^{(1)} = I$$
Así, si suponemos que E es idempotente se tiene que
$$FGFG = FG \Rightarrow F^{(1)}FGFGG^{(1)} = F^{(1)}FGG^{(1)} \Rightarrow GF = I.$$
□

Recordemos que cada transformación lineal de un espacio vectorial de dimensión finita a otro, puede ser representada por una matriz, la cual está determinada de forma única por la transformación lineal y por la elección de las bases de los espacios involucrados.

Una vez seleccionadas las bases, hay una correspondencia uno a uno entre $\mathbb{C}^{m\times n}$, las matrices complejas de $m \times n$, y $\mathcal{L}(\mathbb{C}^n, \mathbb{C}^m)$, el espacio de las transformaciones lineales de $\mathbb{C}^n$ en $\mathbb{C}^m$. Esta correspondencia permite usar el mismo símbolo, por ejemplo A para denotar tanto la transformación lineal $A \in \mathcal{L}(\mathbb{C}^n, \mathbb{C}^m)$ como su representación matricial $A \in \mathbb{C}^{m\times n}$. De esta manera, la ecuación matriz-vector

$$A\mathbf{x} = \mathbf{y} \qquad (A \in \mathbb{C}^{m\times n}, \mathbf{x} \in \mathbb{C}^n, \mathbf{y} \in \mathbb{C}^m)$$

puede ser considerada como que la transformación lineal A envia a $\mathbf{x}$ en $\mathbf{y}$.

En particular, las transformaciones lineales que van de $\mathbb{C}^n$ en $\mathbb{C}^n$ son representadas por las matrices cuadradas de orden n. Profundizando un poco más, el siguiente teorema establece una correspondencia uno a uno entre las matrices idempotentes de orden n y los proyectores $P_{L,M}$ donde $L \oplus M = \mathbb{C}^n$. Más que eso, para cualesquiera dos subespacios complementarios L y M, se obtendrá un método para calcular $P_{L,M}$.

Teorema 3.16. *Para cada matriz idempotente $E \in \mathbb{C}^{n\times n}$, $R(E)$ y $N(E)$ son subespacios complementarios tal que*

$$E = P_{R(E),N(E)}.$$

Recíprocamente, si L y M son subespacios complementarios, hay un único $P_{L,M}$ idempotente tal que $R(P_{L,M}) = L$, $N(P_{L,M}) = M$.

Demostración. Sea E idempotente de orden n. Entonces se sigue de (e) y (g) del Lema 3.14, y de la ecuación

$$\mathbf{x} \;=\; E\mathbf{x} + \mathbf{x} - E\mathbf{x} \;=\; E\mathbf{x} + (I - E)\mathbf{x}, \tag{3.4.1}$$

que $\mathbb{C}^n$ es la suma de $R(E)$ y $N(E)$. Además, $R(E) \cap N(E) = \{0\}$ ya que

$$E\mathbf{x} = (I - E)\mathbf{y} \quad \Rightarrow \quad E\mathbf{x} = E^2\mathbf{x} = E(I - E)\mathbf{y} = O,$$

por (d) del Lema 3.14. Así, $R(E)$ y $N(E)$ son complementarios y (3.4.1) muestra que para cada $\mathbf{x}$, $E\mathbf{x}$ es la proyección de $\mathbf{x}$ en $R(E)$ a lo largo de

$N(E)$, y de esta manera queda establecido que $E = P_{R(E),N(E)}$.
Por otra parte, sean $\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_l\}$ y $\{\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_m\}$ dos bases cualesquiera para L y M, respectivamente. Entonces $P_{L,M}$, si existe, queda determinado de manera única por

$$\begin{cases} P_{L,M}\mathbf{x}_i = \mathbf{x}_i, & (i \in \overline{1,l}), \\ P_{L,M}\mathbf{y}_i = \mathbf{0}, & (i \in \overline{1,m}). \end{cases} \tag{3.4.2}$$

Sea $X = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \dots \quad \mathbf{x}_l]$ la matriz cuyas columnas son los vectores $\mathbf{x}_i$. Similarmente, sea $Y = [\mathbf{y}_1 \quad \mathbf{y}_2 \quad \dots \quad \mathbf{y}_m]$. Entonces (3.4.2) es equivalente a

$$P_{L,M}[X \quad Y] = [X \quad O]. \tag{3.4.3}$$

Como $[X \quad Y]$ es invertible, la única solución de (3.4.3), y por lo tanto de (3.4.2) es

$$P_{L,M} = [X \quad O][X \quad Y]^{-1}. \tag{3.4.4}$$

como (3.4.2) implica que

$$P_{L,M}[X \quad O] = [X \quad O],$$

$P_{L,M}$ expresado como en (3.4.4) es claramente idempotente. $\square$

La relación entre la suma directa en (1.2.2) en la página 3 y el proyector $P_{L,M}$ se da a continuación.

Corolario 3.17. *Sean L y M subespacios complementarios de $\mathbb{C}^n$. Entonces, para cada $x \in \mathbb{C}^n$ la descomposición única de la ecuación (1.2.3) está dada por*

$$P_{L,M}\boldsymbol{x} = \boldsymbol{y}, \qquad (I - P_{L,M})\boldsymbol{x} = \boldsymbol{z}.$$

Si $A^{(1)} \in A\{1\}$, sabemos por el Lema 2.7 que tanto $AA^{(1)}$ como $A^{(1)}A$ son idempotentes, y por tanto, son proyectores. Es de interés saber qué podemos decir acerca de los subespacios asociados a estos proyectores. Ya sabemos por el Corolario 2.8 que

$$R(AA^{(1)}) = R(A), \quad N(AA^{(1)}) = N(A), \quad R((AA^{(1)})^*) = R(A^*). \tag{3.4.5}$$

El siguiente corolario es consecuencia inmediata de estos resultados.

Corolario 3.18. *Si A y X son $\{1, 2\}$-inversas la una de la otra, entonces AX es el proyector en $R(A)$ a lo largo de $N(X)$, y XA es el proyector en $R(X)$ a lo largo de $N(A)$.*

Una aplicación importante de los proyectores es en la clase de matrices diagonalizables (Recordemos que una matriz se dice *diagonalizable* si es similar a una matriz diagonal). Se puede ver que una matriz $A \in \mathbb{C}^{n\times n}$ es diagonalizable si y sólo si tiene n vectores propios linealmente independientes. Esto se usará en la demostración del siguiente teorema, que expresa a una matriz diagonalizable arbitraria como una combinación lineal de proyectores.

Teorema 3.19. (Teorema Espectral para Matrices diagonalizables). *Sea $A \in \mathbb{C}^{n\times n}$ con s distintos valores propios $\lambda_1, \lambda_2, \ldots, \lambda_s$. Entonces A es diagonalizable si y sólo si existen proyectores $E_1, E_2, \ldots, E_s$, tales que*

$$E_iE_j = \delta_{ij}E_i, \tag{3.4.6}$$

$$I_n = \sum_{i=1}^{s} E_i, \tag{3.4.7}$$

$$A = \sum_{i=1}^{s} \lambda_i E_i. \tag{3.4.8}$$

Demostración. $\Leftarrow$] Para $i \in \overline{1,s}$ sea $r_i = rango(E_i)$ y sea $X_i \in \mathbb{C}^{n\times r_i}$ una matriz cuyas columnas son una base para $R(E_i)$. Sea

$$X = [X_1 \quad X_2 \quad \cdots \quad X_s].$$

Entonces por (c) del Lema 3.14, el número de columnas de X es

$$\sum_{i=1}^{s} r_i = \sum_{i=1}^{s} trazaE_i = traza\sum_{i=1}^{s} E_i = trazaI_n = n,$$

por (3.4.7). Por lo tanto X es cuadrada de orden n. Por como se definieron los X_i, existe para cada i un Y_i tal que

$$E_i = X_iY_i$$

Sea

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_s \end{bmatrix}$$

Entonces
$$XY = \sum_{i=1}^{s} X_i Y_i = \sum_{i=1}^{s} E_i = I_n,$$
por (3.4.7). De esta manera X es invertible. Por (e) del Lema 3.14,
$$E_i X_i = X_i,$$
y por lo tanto, por (3.4.6) y (3.4.8),
$$AX = \sum_{i=1}^{s} \lambda_i E_i X_i = [\lambda_1 X_1 \quad \lambda_2 X_2 \quad \cdots \quad \lambda_s X_s] = XD, \tag{3.4.9}$$
donde
$$D = diag(\lambda_1 I_{r_1}, \lambda_2 I_{r_2}, \dots, \lambda_s I_{r_s}). \tag{3.4.10}$$
Como X es invertible se sigue de (3.4.9) que A y D son similares, y con esto A es diagonalizable.
$\Rightarrow$] Si A es diagonalizable
$$AX = XD, \tag{3.4.11}$$
donde X es invertible y D puede ser representada en la forma (3.4.10). Sea X particionada por columnas en $X_1, X_2, \dots, X_s$ conforme a los bloques diagonales de D y, para $i = 1, 2, \dots, s$, sea
$$E_i = [O \quad \cdots \quad O \quad X_i \quad O \quad \cdots \quad O]X^{-1}.$$
En otras palabras, $E_i = \widetilde{X_i} X^{-1}$, donde $\widetilde{X_i}$ denota la matriz obtenida de X reemplazando todas sus columnas, excepto las de X_i, por columnas de ceros. Se verifica fácilmente que E_i es idempotente y que se cumplen (3.4.6) y (3.4.7). Finalmente,
$$\sum_{i=1}^{s} \lambda_i E_i = [\lambda_1 X_1 \quad \lambda_2 X_2 \quad \cdots \quad \lambda_s X_s]X^{-1} = XDX^{-1} = A,$$
por (3.4.11). $\square$

Las matrices idempotentes $\{E_i : i \in \overline{1, s}\}$ son llamadas los idempotentes principales o Frobenius covariantes. La representación de A en (3.4.8) es llamada la descomposición espectral de A.

Note que $R(E_i)$ es el espacio propio de A asociado al valor propio λ_i, mientras que por (3.4.6), $N(E_i)$ es la suma directa de los espacios propios asociados con todos los valores propios de A excepto λ_i.

Un caso especial de los proyectores, son los proyectores ortogonales. Un proyector ortogonal se puede representar, como muestra el siguiente corolario, por una matriz cuadrada que en este caso no es sólo idempotente sino también Hermítica.

Corolario 3.20. *Sea $\mathbb{C}^n = L \oplus M$. Entonces $M = L^\perp$ si y sólo si $P_{L,M}$ es Hermítica.*

Demostración. $\Leftarrow$] Supongamos que $P_{L,M}$ es Hermítica, entonces $P^*_{L,M} = P_{L,M}$. Así, por (1.3.11) del Teorema 1.3

$$M = N(P_{L,M}) = R(P^*_{L,M})^\perp = R(P_{L,M})^\perp = L^\perp.$$

$\Rightarrow$] Supongamos que $M = L^\perp$. Por (a) del Lema 3.14 $P^*_{L,M}$ es idempotente y de esta manera por el Teorema 3.16 es un proyector. Además por (1.3.11) y (1.3.12) del Teorema 1.3

$$\begin{aligned} R(P^*_{L,M}) &= N(P_{L,M})^\perp = M^\perp = L \quad \text{y} \\ N(P^*_{L,M}) &= R(P_{L,M})^\perp = L^\perp = M. \end{aligned}$$

Finalmente, por el Teorema 3.16

$$P^*_{L,M} = P_{L,M}$$

y por lo tanto $P_{L,M}$ es Hermítica. □

Así como hay una biyección entre los proyectores y las matrices idempotentes, el corolario anterior nos muestra que hay una biyección entre los proyectores ortogonales y las matrices Hermíticas idempotentes. En [1] podemos encontrar más propiedades de este tipo de matrices así como el teorema espectral para matrices normales.

3.5. Sistemas Lineales Inconsistentes y Mínimos Cuadrados

Sean $A \in \mathbb{C}^{m\times n}$ y $\mathbf{b} \in \mathbb{C}^m$, el sistema lineal

$$A\mathbf{x} = \mathbf{b} \tag{3.5.1}$$

tiene solución $\mathbf{x}$, si y sólo si $\mathbf{b} \in R(A)$. De lo contrario, el vector residual

$$\mathbf{r} = \mathbf{b} - A\mathbf{x} \tag{3.5.2}$$

es distinto de cero para toda $\mathbf{x} \in \mathbb{C}^n$ y es deseable encontrar una solución aproximada de (3.5.1), que se entenderá como un vector $\mathbf{x}$ que haga al vector residual (3.5.2) lo más "próximo"posible a cero, es decir, que minimice alguna norma de (3.5.2). Una solución aproximada que se usa frecuentemente, especialmente en aplicaciones estadísticas, es la solución de mínimos cuadrados de (3.5.1), definida como el vector $\mathbf{x}$ que minimiza la norma Euclideana del vector residual, es decir, que minimiza la suma de los cuadrados de los módulos de los residuos

$$\sum_{i=1}^{m} |r_i|^2 = \sum_{i=1}^{m} |b_i - \sum_{j=1}^{n} |a_{ij}x_j|^2 = \|\mathbf{b} - A\mathbf{x}\|^2. \tag{3.5.3}$$

El siguiente teorema muestra que $\|A\mathbf{x} - \mathbf{b}\|$ se minimiza escogiendo $\mathbf{x} = X\mathbf{b}$, donde $X \in A\{1,3\}$, estableciendo una relación entre las $\{1,3\}$-inversas y las soluciones de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$, caracterizando así, cada uno de estos dos conceptos en términos del otro.

Teorema 3.21. *Sean $A \in \mathbb{C}^{m\times n}$ y $\boldsymbol{b} \in \mathbb{C}^m$. Entonces $\|A\boldsymbol{x} - \boldsymbol{b}\|$ es mínimo cuando $\boldsymbol{x} = A^{(1,3)}\boldsymbol{b}$, donde $A^{(1,3)} \in A\{1,3\}$. Recíprocamente, si $X \in \mathbb{C}^{n\times m}$ tiene la propiedad de que para todo $\boldsymbol{b}$, $\|A\boldsymbol{x} - \boldsymbol{b}\|$ es mínimo cuando $\boldsymbol{x} = X\boldsymbol{b}$, entonces $X \in A\{1,3\}$.*

Demostración. Podemos ver a $\mathbf{b}$ como

$$\mathbf{b} = (P_{R(A)} + P_{R(A)^\perp})\mathbf{b},$$

así por (1.3.11)

$$\mathbf{b} - A\mathbf{x} = (P_{R(A)}\mathbf{b} - A\mathbf{x}) + P_{N(A^*)}\mathbf{b},$$

donde claramente $P_{R(A)}\mathbf{b} - A\mathbf{x}$ y $P_{N(A^*)}\mathbf{b}$ son ortogonales, por lo tanto por el Lema 1.1

$$\|A\mathbf{x} - \mathbf{b}\|^2 = \|A\mathbf{x} - P_{R(A)}\mathbf{b}\|^2 + \|P_{N(A^*)}\mathbf{b}\|^2. \tag{3.5.4}$$

Aquí observamos que (3.5.4) alcanza su valor mínimo si y sólo si

$$A\mathbf{x} = P_{R(A)\mathbf{b}}. \tag{3.5.5}$$

Afirmamos que esta igualdad se satisface si $\mathbf{x} = A^{(1,3)}\mathbf{b}$ para cualquier $A^{(1,3)} \in A\{1,3\}$. En efecto, por (f) del Lema 2.7, $AA^{(1,3)}$ es idempotente. Luego, por el Teorema 3.16

$$AA^{(1,3)} = P_{R(AA^{(1,3)}),N(AA^{(1,3)})},$$

además por definición $AA^{(1,3)}$ es Hermítica, entonces por el Corolario 3.20 $N(AA^{(1,3)}) = R(AA^{(1,3)})^{\perp}$. Finalmente, por (3.4.5), $R(AA^{(1,3)}) = R(A)$, por lo que $AA^{(1,3)} = P_{R(A)}$. Por lo tanto si $\mathbf{x} = A^{(1,3)}\mathbf{b}$, se tiene la igualdad en (3.5.5).
Recíprocamente, si X es tal que para toda $\mathbf{b}$, $\|A\mathbf{x} - \mathbf{b}\|$ es mínimo cuando $\mathbf{x} = X\mathbf{b}$, (3.5.5) nos da $AX\mathbf{b} = P_{R(A)}\mathbf{b}$ para toda $\mathbf{b}$, y entonces

$$AX = P_{R(A)}.$$

Luego, por el Teorema 3.6, $X \in A\{1,3\}$. □

Corolario 3.22. *Un vector x es solución de mínimos cuadrados para $Ax = b$ si y sólo si*

$$A\boldsymbol{x} = P_R(A)\boldsymbol{b} = AA^{(1,3)}\boldsymbol{b}.$$

Entonces, la solución general de mínimos cuadrados es

$$x = A^{(1,3)}b + (I_n - A^{(1,3)}A)y, \tag{3.5.6}$$

donde $A^{(1,3)} \in A\{1,3\}$ y $y \in \mathbb{C}^n$ arbitrario.

La solución de mínimos cuadrados es única sólo cuando A tiene rango completo de columna. De lo contrario, (3.5.6) es un conjunto infinito de soluciones.

Conclusiones

En el primer capítulo de este trabajo nos pusimos de acuerdo con la notación y se presentaron los conceptos básicos y algunos resultados conocidos del álgebra lineal que se usan en los demás capítulos.

En el segundo capítulo definimos de tres maneras a la inversa de Moore-Penrose y probamos que todas ellas son equivalentes. Posteriormente, las ecuaciones de Penrose que arroja su definición fueron punto de partida para estudiar la existencia y construcción de las inversas generalizadas que cumplen con algunas de estas ecuaciones. En la última sección se expresa a la inversa de Moore-Penrose en términos de conjugadas transpuestas y $\{1\}$-inversas.

El tercer capítulo inicia con una aplicación de las $\{1\}$-inversas en la solución de sistemas lineales; la cual nos dice bajo qué condiciones un sistema lineal tiene solución. Después se caracterizaron los conjuntos $A\{1,3\}$ y $A\{1,4\}$, y algunos subconjuntos de $A\{2\}$. Además, Se estudiaron algunas propiedades de las matrices idempotentes y se probó que existe una relación biunívoca entre éstas y los proyectores. Se presentó también el teorema espectral para matrices diagonalizables; el cual nos dice que podemos ver a una matriz diagonalizable como una combinación lineal de proyectores. Se observó que existe también una relación biunívoca entre los proyectores ortogonales y las matrices Hermíticas idempotentes. Finalmente, se presenta una aplicación de las $\{1, 3\}$-inversas en la solución de mínimos cuadrados.

En un trabajo posterior se pueden analizar con más detalle aplicaciones de las inversas generalizadas, como en la solución de sistemas no lineales, cadenas de Markov finitas, etc., así como estudiar los inversos generalizados de operadores lineales entre espacios de Hilbert.

Bibliografía

[1] Ben-Israel A. y Greville T., *Generalized inverses: theory and applications*, Segunda edición, Springer-Verlag, Nueva York, 2003.

[2] Campbell S.L. y Meyer C.D., *Generalized Inverses of Linear Transformations*, Primera edición, Dover, Nueva York, 1979.

[3] Cruz C. R. y Djordjevic S., *Caracterización de los Inversos Generalizados y Sistemas Lineales*, La Ciencia en tus Manos IX, BUAP, 2008.

[4] Cruz C. R. y Djordjevic S., *Los inversos generalizados de la matriz*, Jóvenes Investigadores, BUAP, 2008.

[5] Djordjevic S. y García B. S. A., *Matrices Idempotentes y Proyecciones de los Inversos Generalizados*, Jóvenes Investigadores Otoño, BUAP, 2009.

[6] Friedberg S.H., Insel A.J. y Spence L.E., *Linear Algebra*, Cuarta edición, Springer-Verlag, Nueva York, 2003.

[7] Lange S., *Linear Algebra*, Tercera edición, Springer-Verlag, Nueva York, 1987.

[8] Takane Y., Takeuchi K. y Yanai H., *Projection Matrices, Generalized Inverse Matrices, and Singular Value Decomposition*, Primera edición Springer Science+Business Media LCC, 2011.

Anotaciones